



Une nouvelle classe de modèles auto-régressifs à valeurs entières

Maher Kachour

► To cite this version:

Maher Kachour. Une nouvelle classe de modèles auto-régressifs à valeurs entières. Mathématiques [math]. Université Rennes 1, 2009. Français. NNT: . tel-00442146

HAL Id: tel-00442146

<https://theses.hal.science/tel-00442146>

Submitted on 18 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE / UNIVERSITÉ DE RENNES 1

sous le sceau de l'Université Européenne de Bretagne

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Mathématiques et Applications

École doctorale MATISSE

présentée par

Maher KACHOUR

préparée à l'UMR 6625 CNRS - IRMAR

Institut de Recherche Mathématique de Rennes

U.F.R. de Mathématique

Intitulé de la thèse

Thèse soutenue à Rennes
le 9 Décembre 2009

devant le jury composé de :

**Une nouvelle classe
de modèles
autorégressifs
à valeurs entières**

Jean-Marc BARDET Professeur
Université de Paris 1 Examineur

Dominique DEHAY Professeur
Université de Rennes 2 Examineur

Bernard DELYON Professeur
Université de Rennes 1 Examineur

Christian FRANCO Professeur
Université de Lille 3 Rapporteur

Anne PHILIPPE Professeur
Université de Nantes Examineur

Jian-Feng YAO Professeur
Université de Rennes 1 Directeur de thèse

A la mémoire de Waël Diab

*C'est avec la logique que nous prouvons
et avec l'intuition que nous trouvons.*
Henri Poincaré

Remerciements

Je tiens à remercier en tout premier lieu le Professeur Jian-Feng Yao, mon directeur de thèse, qui a su me laisser la liberté nécessaire à l'accomplissement de mes travaux, tout en gardant un œil critique et avisé. La confiance que vous m'avez accordée ainsi que nos nombreuses discussions m'ont permis de progresser et de mieux appréhender les différentes facettes du métier d'enseignant-chercheur.

Mes remerciements vont également à F. Jay Breidt et Christian Francq, je suis très touché de l'honneur que vous me faites en acceptant de juger ce travail et d'en être les rapporteurs. Veuillez accepter l'expression de ma plus profonde gratitude.

Je présente toute ma reconnaissance à Anne Philippe, Jean-Marc Bardet, Dominique Dehay et Bernard Delyon, pour avoir accepté d'examiner ma dissertation doctorale et de faire partie de mon jury de thèse. Soyez assurés, Madame, Messieurs, de toute mon estime et de mon profond respect.

Je voudrais adresser un remerciement à tous les personnels de l'UFR de Mathématique et de l'IRMAR, pour leur compétence, leur écoute, leur gentillesse et leur disponibilité. Je pense particulièrement à Claude Boschet (maman Claude), Marie-Aude Verger, Hélène Rousseaux, Karine Falc'Hon, Chantal Halet, Marie-Annick Paulmier, Claudine Helies, Dominique Hervé et Patrick Perez.

Je ne sais comment exprimer ma gratitude à tous les enseignants avec qui j'ai collaboré pendant ces trois années. Je vous remercie chaleureusement pour votre confiance. Je souhaite évidemment remercier toutes les personnes rencontrées, lors des manifestations scientifiques, pour l'intérêt qu'elles ont porté à mon travail.

La réalisation de cette thèse fut aussi l'occasion d'une aventure humaine merveilleuse. Au cours de ces trois dernières années, j'ai croisé la route de nombreuses personnes. Je reconnais que chacune a, à des degrés divers, apporté une contribution positive à ma vie personnelle et professionnelle. Merci à Adrien Saumard (mon poto), je n'oublierai jamais

nos discussions, nos nuits blanches et nos fous rires. Je tiens à remercier également Matthieu et Marianne ainsi que tous les membres de la famille Saumard, votre hospitalité me fait chaud au cœur.

"Les amis sont la famille que nous choisissons" et au sein de l'IRMAR j'ai trouvé une grande famille. Ainsi, mes plus vifs remerciements s'adressent à Christian Naumovic (naumov), Frédérique le Louër (fred), Nirmal Antonio Tamarasselvame (nini), Arnaud Jobin (jobinho), Nicolas Bideau (nico), Solena Bravo-Pinto (tata), Victor Péron (toto), Jonathan Marco (johnny), Ludovic Goudenège (le p'tit ludo), Adrien Richou (le fils caché de B.D.), Thomas Sierocinski, Audrey Houillier (les gens du nord), Alina Crudu (chou-chou), Polynice Eyi (poly), Rodolphe Richard (coffee-man), Jimmy Lamboley (salsa-man), Anne-Claire Bennis (cacahuète), Colas Bardavid (mec à la cool), et Géraldine Pichot (gerald). Merci les ami(e)s pour votre écoute, vos encouragements, vos conseils et surtout pour tous les moments de bonheur partagés.

Je remercie de plus tous ceux qui sont passés un jour prendre un café au bureau 434, ceux que je n'ai pas eu la chance de côtoyer plus souvent, et ceux avec qui j'ai partagé le repas de midi au restaurant universitaire. Merci à tous les autres doctorants pour le monde parallèle qui s'est créé autour de leur jeu de hasard. Grâce à vous, j'ai réalisé que l'ouverture c'est comprendre la divergence des points de vue.

Je salue aussi les anciens et les nouveaux doctorants. Ainsi, je tiens à exprimer ma profonde gratitude à toutes les personnes que je n'ai pas citées ici et qui se reconnaîtront dans ces quelques lignes.

Loin de la tour des maths, j'exprime toute mon amitié à Bilal Nehmeh, Mohamad Muhieddine, Ayman Khalil, Hassan Sabbah, Hussein Sabbah, Carole Cabreton, Benoît Patra, et Tassadit Abid. Je m'excuse auprès de toutes les personnes que j'apprécie et dont le nom n'apparaît pas dans cette page. Je ne saurais pas les citer toutes sans dépasser le nombre de pages raisonnablement admis dans ce genre de travail.

Une pensée émue pour tous les étudiants avec qui j'ai partagé une salle pendant ces trois années. Merci pour votre confiance et votre soutien.

Cela va de soi, je remercie évidemment ma famille pour son irremplaçable et inconditionnel soutien. Ils ont été présents, malgré la distance, pour écarter les doutes, soigner les blessures et partager les joies. Il m'est impossible d'exprimer toute ma reconnaissance et mon respect pour leurs sacrifices. Cette thèse est aussi un peu la leur.

En vérité, le chemin importe peu, la volonté d'arriver suffit à tout.

Albert Camus

Préface

Les séries chronologiques à valeurs entières sont fréquemment utilisées dans la pratique. Citons quelques exemples : nombre des transactions hebdomadaires d'un titre boursier, la température quotidienne à Rennes pendant cet été, le calibrage d'un fruit donné, et bien d'autres.

Avant la fin des années 70 de telles observations étaient traitées par des modèles réels (par exemple, ARMA, VAR, ARCH,...) bien connus dans la littérature, sans tenir compte de la nature entière de ces observations.

Une première tentative de modélisation des séries chronologiques à valeurs entières est due à Jacobs & Lewis (1978). Leur modèle peut être décrit comme une combinaison linéaire aléatoire de variables aléatoires i.i.d. discrètes.

Plus tard, des auteurs tels McKenzie (1985), Al-Osh & Alzaid (1987) et bien d'autres ont introduit des classes de modèles qui possèdent les mêmes structures de corrélation que des modèles réels, tout en respectant la particularité entière des séries observées.

En effet, pour construire tels modèles les auteurs ont remplacé la multiplication usuelle dans les modèles réels (par exemple dans un modèle $AR(p)$) par un opérateur d'amincissement (par exemple l'opérateur d'amincissement binomial de Steutel & Van Harn (1979)).

Cependant, ce remplacement fait apparaître plusieurs contraintes limitant l'utilisation de résultats déjà connus pour les modèles réels.

Au cours de cette thèse, notre intérêt s'est tourné uniquement vers la recherche des nouvelles classes de modèles autorégressifs à valeurs entières. Dans ce but, nous nous sommes basés sur l'opérateur d'arrondi à l'entier le plus près, pour construire le modèle $RINAR(p)$ (pour p -order rounded integer-valued autoregressif). L'idée de la construction de ce dernier est simple, il suffit de censurer (en utilisant l'opérateur d'arrondi) la fonction de régression linéaire d'un processus $AR(p)$ ayant comme bruit une suite de variables aléatoires centrées i.i.d. à valeurs dans \mathbb{Z} .

Ainsi, par construction les modèles proposés exploitent un domaine rarement examiné : la modélisation des séries chronologiques à valeurs entières ayant des observations négatives. Notons qu'en général les processus basés sur les opérateurs d'amincissement sont appropriés pour la modélisation des séries chronologiques à valeurs entières positives, en particulier les séries de comptage. Par ailleurs, comparés aux modèles existants dans la littérature ces nouveaux modèles possèdent plusieurs avantages.

Après une brève présentation de quelques modèles autorégressifs à valeurs entières basés sur les opérateurs d'amincissement, nous suscitons leurs limitations, et nous argumentons l'introduction de la nouvelle classe de modèles basés sur l'opérateur d'arrondi (Chapitre 1). Nous traitons, dans le Chapitre 2, le modèle RINAR du premier ordre. Le modèle RINAR d'ordre supérieur est examiné dans le Chapitre 3. Dans le Chapitre 4, nous étudions le modèle RINAR vectoriel approprié pour analyser les séries chronologiques à valeurs entières multivariées. Le Chapitre 5 est consacré à l'étude du modèle RINAR centré. Finalement, dans le Chapitre 6, nous présentons le modèle PRINAR dédié à l'analyse des séries chronologiques à valeurs entières positives.

*Il est toujours aisé d'être logique. Il est presque impossible
d'être logique jusqu'au bout.*

Albert Camus

Table des matières

1	Introduction	1
1.1	Les opérateurs d’amincissement	3
1.2	Quelques modèles basés sur les opérateurs d’amincissement	6
1.2.1	Le modèle INAR(1)	6
1.2.2	Le modèle INAR(p)-AA	9
1.2.3	Le modèle INAR(p)-DL	10
1.2.4	Le modèle GINAR(p)	11
1.2.5	Le modèle MGINAR(p)	12
1.3	Contributions de la thèse	13
1.3.1	Critiques des modèles basés sur les opérateurs d’amincissement . . .	13
1.3.2	Intérêts des modèles introduits	14
1.3.3	Plan de la thèse	16
2	Le modèle RINAR(1)	19
2.1	Quelques définitions	20
2.2	Stationnarité et ergodicité du processus RINAR(1)	24
2.3	Propriétés du processus RINAR(1)	26
2.4	Estimation des paramètres	29
2.4.1	La consistance forte de l’estimateur des moindres carrés quand le coefficient de régression α_0 est irrationnel	34
2.4.2	La consistance forte de l’estimateur des moindres carrés quand le coefficient de régression α_0 est rationnel	43
2.5	Méthode numérique pour calculer $\hat{\theta}_n$	57
2.5.1	Initialisation	58
2.5.2	Recherche dichotomique successive	58

2.5.3	Étude de simulation	61
2.6	Analyse des données d'O'Donovan	65
2.6.1	Le modèle AR(1) d'O'Donovan	67
2.6.2	Ajustement à un modèle RINAR(1)	69
2.6.3	Comparaison et commentaires	71
2.7	Analyse des taux de variation annuels de la population Suédoise	71
2.7.1	Ajustement à un modèle AR(1)	74
2.7.2	Ajustement à un modèle RINAR(1)	74
2.7.3	L'approche de McCleary & Hay	76
2.7.4	Comparaison et commentaires	77
3	Le modèle RINAR(p)	79
3.1	Stationnarité et ergodicité du modèle RINAR(p)	80
3.2	Estimation des paramètres	84
3.2.1	La consistance forte de l'estimateur des moindres carrés quand un au moins des coefficients de régression α_j^* est irrationnel	91
3.2.2	La consistance forte de l'estimateur des moindres carrés quand tous coefficients de régression α_j^* sont rationnels	95
3.3	Méthode numérique pour calculer $\hat{\theta}_n$	104
3.3.1	Initialisation	104
3.3.2	Recherche dichotomique successive	104
3.3.3	Une expérience de simulation	106
3.4	Analyse des données de Fürth	110
3.4.1	Le modèle INAR(2) de Jung & Termayne	111
3.4.2	Ajustement à un modèle RINAR(2)	112
4	Le modèle RINVAR(1)	117
4.1	Stationnarité et ergodicité du modèle RINVAR(1)	118
4.2	Estimation des paramètres	122
4.2.1	La consistance forte de l'estimateur des moindres carrés quand chaque ligne de la matrice M_0 possède un coefficient irrationnel	127
4.2.2	La consistance forte de l'estimateur des moindres carrés quand tous les coefficients de la matrice M_0 sont rationnels	130

4.2.3	La consistance forte de l'estimateur des moindres carrés quand au moins une ligne de la matrice M_0 possède au moins un coefficient irrationnel	134
4.3	Méthode numérique pour calculer $\hat{\theta}_n$	137
4.4	Analyse des taux de variation et les indices de récolte annuels de la population Suédoise	138
4.4.1	L'approche de McCleary & Hay	139
4.4.2	Ajustement à un modèle RINVAR(1)	140
4.4.3	Comparaison et commentaires	142
5	Le modèle RINAR(1) centré	145
5.1	Quelques définitions	146
5.2	Étude du modèle RINAR(1) centré	148
5.2.1	Stationnarité et ergodicité du processus RINAR(1) centré	149
5.2.2	Estimation du paramètre α	149
5.3	Propriétés du processus RINAR(1) centré	150
5.3.1	Comparaison avec un AR(1) centré	153
5.3.2	Étude de simulation	155
5.4	Méthode numérique pour calculer $\hat{\alpha}_n$	160
5.5	Le modèle RINAR(1) recentré	163
6	Le modèle PRINAR(1)	165
6.1	Introduction	166
6.2	Étude théorique du processus PRINAR(1)	167
6.3	Propriétés du processus PRINAR(1)	168
6.3.1	Le processus PRINAR(1) recentré	171
6.3.2	Étude de simulation	172
6.4	Support de la loi stationnaire	177
	Conclusion et perspectives	179
	Table des figures	185
	Bibliographie	189

Chapitre 1

Introduction

Sommaire

1.1	Les opérateurs d'amincissement	3
1.2	Quelques modèles basés sur les opérateurs d'amincissement . .	6
1.2.1	Le modèle INAR(1)	6
1.2.2	Le modèle INAR(p)-AA	9
1.2.3	Le modèle INAR(p)-DL	10
1.2.4	Le modèle GINAR(p)	11
1.2.5	Le modèle MGINAR(p)	12
1.3	Contributions de la thèse	13
1.3.1	Critiques des modèles basés sur les opérateurs d'amincissement .	13
1.3.2	Intérêts des modèles introduits	14
1.3.3	Plan de la thèse	16

Dans certaines situations il devient nécessaire de traiter les séries chronologiques à valeurs entières. Ces dernières peuvent se produire dans beaucoup de domaines, par exemple, dans la théorie d'assurance, la médecine, les systèmes de files d'attente, les communications, la fiabilité, la météorologie, et bien d'autres. Les processus de comptage sont des exemples de ces séries chronologiques et ce qu'ils comptent diffère d'un processus à l'autre. Ceux-ci peuvent être les nombres d'accidents, patients, victimes de crime, messages transmis, erreurs détectées, et ainsi de suite (voir par exemple, Johansson (1996, [36]), Freeland (1998, [25]), Ye et al. (2001, [65]), Pokropp et al (2006, [56]), Lambert and Liu (2006, [45]), Jung et al. (2006, [37]) et Weiß(2008, [62])).

Au premier regard, l'analyse de telles séries chronologiques peut présenter quelques difficultés. Si l'analyse est basée sur quelques modèles stochastiques, les modèles doivent

refléter la particularité entière de la série observée. Pour cette raison, nous préférons ne pas considérer, à titre d'exemple, un processus autorégressif à valeurs réelles, pour analyser de telles observations. Rappelons que les modèles classiques tels ARMA, ARCH et d'autres, bien connus dans la littérature, génèrent des observations réelles (voir Brockwell et Davis (2002, [12])).

La recherche des modèles stochastiques pour des observations à valeurs entières a débuté à la fin des années 1970. Jacobs & Lewis (1978 et 1983, [32] [33] et [34]) ont introduit la première classe de modèles, notée DARMA. Cette dernière peut être décrite comme une combinaison linéaire aléatoire de variables aléatoires i.i.d. discrètes. L'approche de Jacobs-Lewis peut être considérée comme analogue à celle de Box-Jenkins dans laquelle la combinaison linéaire des valeurs réelles est remplacée par un mélange probabiliste. Les processus DARMA permettent de traiter les processus prenant un nombre fini de valeurs et, plus particulièrement, les séries binaires.

Les premiers processus stochastiques à valeurs entières dont les propriétés sont similaires à celles des modèles réels classiques sont dus à McKenzie (1985, [50]) et Al-Osh & Alzaid (1987, 1990, [2], [6]). Ces modèles sont basés sur l'opérateur de Steutel & Van Harn (1979, [59]). Dans Al-Osh & Alzaid (1990, [6]), le processus INAR(p) (pour processus AR(p) à valeurs entières) possède une structure de corrélation similaire à celle du processus ARMA(p, p-1). Du & Li (1991, [21]) ont suggéré une autre spécification du modèle INAR(p) dont la structure de corrélation est identique à celle du processus AR(p) standard. Gauthier & Latour (1994, [30]) et Latour (1998, [47]) ont considéré une version plus générale du modèle, noté GINAR(p), basée sur l'opérateur généralisé de Steutel & Van Harn. Latour (1997, [46]) a introduit le modèle MGINAR(p), basé sur l'opérateur matriciel généralisé de Steutel & Van Harn, pour les séries chronologiques à valeurs entières multivariées.

Les modèles à moyenne mobile à valeurs entières, notés INMA(q), sont basés aussi sur l'opérateur de Steutel & Van Harn. Le processus INMA a été étudié par McKenzie (1988, [52]). Al-Osh & Alzaid (1988, 1991, [3], [4]) ont proposé une autre spécification. Pour une comparaison entre les deux spécifications, nous nous référons à Brännäs et Hall (2001, [9]). Pour une revue récente sur les modèles de série chronologique à valeurs entières possédant la même structure de corrélation que la classe ARMA, nous nous référons à Kedem et Fokianos (2002, [44]), McKenzie (2003, [53]) et Jung et Termayne (2006, [40]).

Par ailleurs, nous mentionnons les récents travaux de Ferland et al. (2006, [24]) concernant la classe de modèles GARCH à valeurs entières, et Doukhan et al. (2006, [19]) où ils

traitent le modèle bilinéaire simple à valeurs entières. Ces modèles sont basés sur l'opérateur d'aminicissement généralisé.

Dans ce chapitre, nous introduisons les opérateurs d'aminicissement, en particulier l'opérateur d'aminicissement binomial de Steutel & Van Harn (1979, [59]). À la suite, nous énonçons quelques classes de modèles autorégressifs à valeurs entières, basés sur les opérateurs d'aminicissement : notamment, les modèles INAR(p), les modèles GINAR(p) et les modèles MGINAR(p). En conséquent, nous citons les limites de ces modèles et nous présentons les motivations et les intérêts d'introduire une nouvelle classe de modèles autorégressifs à valeurs entières, basés sur l'opérateur d'arrondi à l'entier le plus près.

Finalement, nous donnons le plan de la thèse.

1.1 Les opérateurs d'aminicissement

Pour obtenir une structure de corrélation analogue à celle de la classe ARMA, beaucoup de modèles utilisent des opérations d'aminicissement.

L'opérateur d'aminicissement (thinning operator) le plus connu est celui de Steutel & Van Harn (1979, [59]). Ce dernier est appelé l'opérateur d'aminicissement binomial.

Définition 1.1.1. (*L'opérateur \circ de Steutel & Van Harn*). Soit (ξ_k) une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètre $a \in [0, 1]$ et indépendante de Z , une variable aléatoire à valeurs entières positives.

L'opérateur d'aminicissement, noté \circ , est défini par

$$a \circ Z = \sum_{k=1}^Z \xi_k.$$

Nous appelons la suite (ξ_k) une série de comptage. Par conséquent, conditionnellement à Z la variable aléatoire $a \circ Z$ suit une loi binomial de paramètres Z et a .

Pour l'interprétation de l'opérateur d'aminicissement binomial, considérez une population de taille Z , à un moment t . Si nous observons la même population à un moment postérieur, dit $t + 1$, alors la population peut être diminuée, parce que certains individus sont morts entre les moments t et $t + 1$. Si les individus meurent indépendamment les uns des autres, et si la probabilité de mourir entre t et $t + 1$ est égale à $1 - a$ pour tous les individus, alors le nombre des survivants est donné par $a \circ Z$. Maintenant, nous donnons quelques propriétés

de base de l'opérateur d'amincissement binomial.

$$\begin{cases} \mathbb{E}[a \circ Z] = a \mathbb{E}[Z], \\ V[a \circ Z] = a^2 V[Z] + a(1-a) \mathbb{E}[Z], \\ \text{cov}[a \circ Z, Z] = a V[Z]. \end{cases} \quad (1.1)$$

Pour une preuve explicite de ces dernières et bien d'autres propriétés nous nous référons à Freeland (1998, [25]) et da Silva (2005, [14]).

En fait, le concept d'amincissement binomial a été introduit pour adapter les termes de la classe des distributions auto-décomposables discrètes, notée DSD (discrete self-decomposable). Rappelons qu'une distribution sur \mathbb{R} est dite auto-décomposable ou de classe L si sa fonction caractéristique φ satisfait (voir par exemple, Lukacs (1970, [48]), page 161)

$$\varphi(t) = \varphi(\alpha t) \varphi_\alpha(t), \quad t \in \mathbb{R} \text{ et } \alpha \in [0, 1], \quad (1.2)$$

où φ_α est une fonction caractéristique. Ceci implique que, pour les variables aléatoires correspondantes, nous avons

$$X \stackrel{d}{=} \alpha X' + X_\alpha \text{ avec } \alpha \in [0, 1], \quad (1.3)$$

où X' et X_α sont indépendantes et X' a la même distribution que X .

Définition 1.1.2. Une distribution sur \mathbb{N} avec une fonction génératrice P est dite DSD si

$$P(z) = P(1 - a + az)P_a(z) \quad |z| \leq 1; \quad a \in [0, 1], \quad (1.4)$$

avec P_a une fonction génératrice.

Définition 1.1.3. Une variable aléatoire Z à valeurs dans \mathbb{N} est dite DSD si pour tout $a \in [0, 1]$, il existe une variable aléatoire discrète ε_a , indépendante de $a \circ Z$, telle que

$$Z \stackrel{d}{=} a \circ Z + \varepsilon_a. \quad (1.5)$$

Nous pouvons mentionner ici que la classe DSD est une sous-classe assez large de la classe des distributions discrètes infiniment divisible définie comme suit.

Définition 1.1.4. Soit Z une variable aléatoire à valeurs dans \mathbb{N} . Sa fonction génératrice $P_X(z)$ est dite infiniment divisible si $(P_X(z))^{\frac{1}{n}}$ est elle-même une fonction génératrice pour tout $n \in \mathbb{N}^*$.

Feller (1968, [23]) a montré qu'une distribution sur \mathbb{N} est infiniment divisible, si et seulement si elle est une distribution de Poisson composée.

Notons que beaucoup de distributions, y compris binomial Négative, Poisson et Poisson généralisée (les propriétés de cette distribution ont été étudiée par Consul (1989, [13]) et Ambagaspitiya et Balakrishnan (1994, [8])), appartiennent à la classe des distributions DSD (voir Zhu and Joe (2003, [67])). En particulier, cette dernière contient la classe des distributions discrètes stables et tous ces éléments sont unimodales (voir Steutel & Van Harn (1979, [59])).

Pour adapter le concept d'amincissement binomial à différents types d'interprétations ou des processus, quelques modifications concernant les hypothèses de la Définition 1.1.1 ont été proposées. Brännäs and Hellström (2001, [10]) ont proposé de tenir compte de la dépendance entre les indicateurs de la série de comptage (ξ_k) de la Définition 1.1.1, puisque ces indicateurs peuvent représenter des individus dans le même milieu macro-économique et donc affectés de la même façon.

Une autre généralisation de l'opérateur d'amincissement binomial a été proposée par Lattour (1998, [47]). Ici, les éléments de la série de comptage ξ_k de la Définition 1.1.1 sont des variables aléatoires à valeurs dans \mathbb{N} de moyenne $a \in [0, 1]$ et variance finie λ (voir Section 1.2.4). Cette opération est connue sous le nom d'amincissement généralisé. Elle peut être interprétée comme un processus de reproduction qui permet aux individus de se reproduire plus d'une fois (comme dans le cas d'amincissement binomial) : les ξ_k désigne le nombre de descendants d'une certaine personne, y compris éventuellement cet individu lui-même. L'opérateur d'amincissement généralisé comprend l'amincissement binomial comme cas particulier ($\lambda = a(1-a)$), ainsi que l'opérateur d'amincissement étendu (extended thinning) proposé par Zhu and Joe (2003, [67]). Notons qu'en utilisant l'opérateur d'amincissement étendu, les auteurs (2003, [67]) ont défini la classe des distributions discrètes auto-décomposable généralisée, notée $\text{DSD}(\gamma)$ où $\gamma \in [0, 1[$. Notons que, les classes $\text{DSD}(0)$ et DSD ont les mêmes propriétés et tous les éléments de la classe $\text{DSD}(\gamma)$ sont infiniment divisibles. De plus, l'opérateur d'amincissement généralisé possède les mêmes propriétés que l'opérateur d'amincissement binomial.

D'autre part, Joe (1996, [35]) et Zheng et al. (2007, [66]) ont proposé d'étendre le concept d'amincissement en permettant au paramètre a de la Définition 1.1.1 d'être aléatoire lui-même. Ainsi, l'opération d'amincissement résultante est dite d'amincissement à coefficient aléatoire (random coefficient thinning).

1.2 Quelques modèles basés sur les opérateurs d'amincissement

La première classe de modèles autorégressifs à valeurs entières, notée INAR, a été introduite par McKenzie (1985, [50]) et Al-Osh & Alzaid (1987, 1990, [2], [6]).

Définition 1.2.1. Soit $\{X_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires à valeurs entières positives ; $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires i.i.d. à valeurs entières positives $p \in \mathbb{N}^*$ et $\{a_j\}_{j \in \{1,2,\dots,p\}}$, une suite de constantes telles que $\forall j, 0 \leq a_j < 1$ ($a_p > 0$). Alors, $\{X_t\}_{t \in \mathbb{Z}}$ est un processus INAR(p) si

$$X_t = \sum_{j=1}^p a_j \circ X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}. \quad (1.6)$$

1.2.1 Le modèle INAR(1)

Le modèle INAR(1) a été étudié par Al-Osh & Alzaid (1987, [2]). Ainsi, ce dernier est défini par

$$X_t = a \circ X_{t-1} + \varepsilon_t, \quad \forall t \in \mathbb{Z},$$

où le bruit (ε_t) est une suite de variables aléatoires i.i.d. à valeurs dans \mathbb{N} , avec $\mathbb{E}\varepsilon_t = \lambda$ et $V(\varepsilon_t) = \sigma^2$. Ainsi, nous avons

$$a \circ X_{t-1} = \sum_{k=1}^{X_{t-1}} \xi_k, \quad (1.7)$$

où (ξ_k) est une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètre $a \in [0, 1]$, indépendante de X_{t-1} et (ε_t) est indépendante de (ξ_k) . Par conséquent, $(a \circ X_t \mid X_t)$ suit une loi binomial de paramètres X_t et a .

Nous pouvons considérer une réalisation X_t comme ayant deux composantes : les éléments survivants de X_{t-1} , avec une probabilité de survie a pour chacun, et les éléments qui entrent dans le système, dans l'intervalle $]t-1, t]$ (le terme d'innovation ε_t). Ainsi, le modèle INAR(1) peut être considéré comme un processus de branchement de Galton-Watson avec immigration. Pour une définition et un développement historique de ce dernier, nous nous référons à Heyde et Seneta (1972, [31]). Finalement, le modèle INAR(1) est également lié à la file d'attente $M/M/\infty$ (voir McKenzie (1988, [52])).

Avec ces interprétations, le modèle INAR(1) s'applique à de nombreuses situations dans la pratique. À titre d'exemple, X_t peut décrire le nombre de clients. ε_t décrit les nouveaux clients, et $X_{t-1} - a \circ X_{t-1}$ est le nombre de clients qui ont été perdus à la fin de la dernière

période. Brännäs et al. (2002, [11]) ont utilisé cette approche pour modéliser le nombre des clients d'un hôtel.

Sous l'hypothèse $0 \leq a < 1$, la stationnarité d'INAR(1) est assurée.

Les moments du premier et second ordre du modèle sont

$$\mathbb{E}X_t = \frac{\lambda}{1-a} \text{ et } V(X_t) = \frac{\sigma^2 + \lambda a}{1-a^2}. \quad (1.8)$$

De plus, beaucoup de ses propriétés sont similaires à celles du processus AR(1) standard.

En particulier, la fonction d'autocorrélation (ACF) vaut

$$\rho(k) \equiv \text{corr}(X_t, X_{t-k}) = a^k, \quad \forall k \in \mathbb{N}. \quad (1.9)$$

Ainsi, l'estimateur de Yule-Walker pour le paramètre a n'est autre que le coefficient d'autocorrélation empirique de premier ordre, c.à.d.,

$$\hat{a} = \hat{\rho}(1) \quad (1.10)$$

L'estimation de Yule-Walker de λ est basée sur le moment du premier ordre, c.à.d.

$$\hat{\lambda} = \bar{X}(1 - \hat{a}), \text{ où } \bar{X} \text{ est la moyenne empirique.} \quad (1.11)$$

La prévision à un pas à l'instant T , basée sur l'espérance conditionnelle, est linéaire (comme pour un AR(1) réel) et donnée par

$$\hat{X}_{T+1} = \mathbb{E}(X_{T+1} \mid \mathcal{F}_T) = aX_T + \lambda, \text{ où } \mathcal{F}_T = \sigma\{X_T, X_{T-1}, \dots\}. \quad (1.12)$$

Finalement, la probabilité de transition du processus INAR(1) est donnée par

$$\mathbb{P}[X_{t+1} = k \mid X_t = l] = \sum_{j=0}^{\min(k,l)} \binom{l}{j} a^j (1-a)^{l-j} \mathbb{P}(\varepsilon_t = k-j). \quad (1.13)$$

D'autre part, Al-Osh & Alzaid (1988, [5]) ont calculé la fonction génératrice de X_t

$$P_{X_t}(s) = P_{X_0}(1 - a^t + a^t s) \prod_{k=0}^{t-1} P_\varepsilon(1 - a^k - a^k s), \quad |s| \leq 1, \quad (1.14)$$

où P_ε est la fonction génératrice de ε_1 . Puisque $\mathbb{E}[\varepsilon_1] < \infty$, la limite $\lim_{t \rightarrow \infty} P_{X_t}(s)$ existe et vaut

$$P_X(s) = \prod_{k=0}^{\infty} P_\varepsilon(1 - a^k - a^k s), \quad |s| \leq 1. \quad (1.15)$$

Par ailleurs, à partir de la Définition 1.2.1 (pour $p = 1$) il est simple de vérifier que

$$P_X(s) = P_\varepsilon(s)P_X(1 - a + as), \quad (1.16)$$

où X est une variable aléatoire ayant comme distribution la loi stationnaire du processus INAR(1). D'autre part, l'équation (1.16) est liée à la définition de la classe des distributions auto-décomposables discrètes, notée DSD (voir Section 1.1). Ainsi, nous constatons que la distribution marginale du modèle stationnaire INAR(1) est un élément de la classe DSD. Notons que, Al-Osh & Alzaid (1987, [2]) ont présenté la loi marginale du modèle INAR(1) en fonction des termes du bruit (ε_t) ,

$$X_t \stackrel{d}{=} \sum_{j=0}^{\infty} a^j \circ \varepsilon_{t-j}. \quad (1.17)$$

Ainsi, sous l'hypothèse où ε_t suit une loi de Poisson de paramètre λ (c.à.d. $\varepsilon_t \rightsquigarrow Po(\lambda)$, $\mathbb{E}\varepsilon_t = V(\varepsilon_t) = \lambda$), la loi stationnaire de la chaîne de Markov (X_t) est aussi de Poisson de paramètre $\frac{\lambda}{1-a}$. Par conséquent, $\mathbb{E}X_t = V(X_t) = \frac{\lambda}{1-a}$ et ce modèle sera appelé PoINAR(1). Pour d'autres propriétés concernant ce dernier, nous nous référons à Freeland (1998, [25]), Freeland et McCabe (2004, 2005, [26], [27]) et Weiß(2008, [63]).

Une loi de Poisson n'est pas toujours appropriée à la modélisation et l'analyse des séries chronologiques à valeurs entières. C'est parce que la moyenne et la variance de la distribution de Poisson sont égales et que cette propriété n'est pas toujours vérifiée pour des données réelles. Autres distributions marginales possibles pour le processus stationnaire INAR(1) sont la loi binomial Négative (voir par exemple, Al-Osh et Aly (1992, [1]), Zhu et Joe (2006, [68]), et Weiß(2008, [64])), la loi géométrique (voir par exemple, McKenzie (1986, [51]) et Ristić et al. (2008, [58])), et la loi de Poisson généralisée (voir Alzaid et Al-Osh (1993, [7])).

Soit X_1, \dots, X_d une série chronologique à valeurs dans \mathbb{N} , la loi marginale peut être identifiée à l'aide de l'histogramme, et l'ordre du modèle avec la fonction d'autocorrélation empirique. D'ailleurs, Jung et Tremayne (2003, [39]) ont fourni des approches pour examiner la dépendance entre les observations.

Une fois que le modèle INAR(1) a été identifié comme approprié pour la série chronologique X_1, \dots, X_d , l'estimation des paramètres du modèle peut être faite, entre autres, par les trois approches suivantes : la méthode des moments (MM), les moindres carrés conditionnel (MCC), et le maximum de vraisemblance (MV). Pour une étude comparative détaillée des propriétés et les performances de ces derniers et d'autres différents types d'estimateurs nous nous référons à Jung et al. (2005, [38]).

1.2.2 Le modèle INAR(p)-AA

Maintenant, nous considérons le processus INAR(2), défini par

$$X_t = a_1 \circ X_{t-1} + a_2 \circ X_{t-2} + \varepsilon_t, \quad \forall t \in \mathbb{Z},$$

où le bruit (ε_t) est une suite de variables aléatoires i.i.d. à valeurs dans \mathbb{N} , avec $\mathbb{E}\varepsilon_t = \lambda$ et $V(\varepsilon_t) = \sigma^2$, $0 \leq a_1 < 1$ et $0 < a_2 < 1$. Par suite, nous avons

$$a_1 \circ X_{t-1} = \sum_{k=1}^{X_{t-1}} \xi_{1k} \text{ et } a_2 \circ X_{t-2} = \sum_{k=1}^{X_{t-2}} \xi_{2k}, \quad (1.18)$$

où la série de comptage (ξ_{1k}) (resp. (ξ_{2k})) est une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètre a_1 (resp. a_2) et indépendante de X_{t-1} (resp. X_{t-2}). De plus, (ε_t) est indépendante de (ξ_{1k}) et (ξ_{2k}) .

Le choix des séries de comptage est crucial pour déterminer la structure de corrélation du processus. Ainsi, nous distinguons deux différentes spécifications.

Le modèle INAR-AA, a été introduit par Al-Osh & Alzaid (1990, [6]).

Ce dernier est une extension directe du processus INAR(1). Pour assurer la stationnarité du modèle les auteurs supposent que $a_1 + a_2 < 1$. Ici, conditionnellement à X_t le vecteur $(a_1 \circ X_t, a_2 \circ X_t, X_t - a_1 \circ X_t - a_2 \circ X_t)$ suit une loi Trinomiale de paramètre $(X_t; a_1, a_2, 1 - a_1 - a_2)$. C'est une extension naturelle multivariée de l'hypothèse équivalente dans le processus INAR(1) où $(a \circ X_t \mid X_t) \rightsquigarrow B(X_t, a)$.

Il en découle que le modèle INAR(p)-AA possède une structure de corrélation similaire à celle du processus ARMA(p, p-1). Sous l'hypothèse où ε_t suit une loi de Poisson de paramètre λ , la loi stationnaire de la chaîne de Markov (X_t) est aussi de Poisson de paramètre $\frac{\lambda}{1 - a_1 - a_2}$. Par conséquent, $\mathbb{E}X_t = V(X_t) = \frac{\lambda}{1 - a_1 - a_2}$, ce modèle sera appelé PoINAR(2)-AA. L'ACF du processus est donnée par

$$\rho(1) = a_1 \text{ et } \rho(k) = a_1 \rho(k-1) + a_2 \rho(k-2), \text{ pour } k \geq 2. \quad (1.19)$$

Alors, les estimateurs de Yule-Walker pour les paramètres a_1 et a_2 sont

$$\hat{a}_1 = \hat{\rho}(1) \text{ et } \hat{a}_2 = \hat{\rho}(2) - \hat{\rho}(1)^2. \quad (1.20)$$

Par conséquent, l'estimation de Yule-Walker de λ est

$$\hat{\lambda} = \bar{X}(1 - \hat{a}_1 - \hat{a}_2), \text{ où } \bar{X} \text{ est la moyenne empirique.} \quad (1.21)$$

La fonction de la moyenne conditionnelle (régression) de ce modèle est non-linéaire.

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = \lambda \left[1 + (a_1 + Ua_1a_2) \frac{p(y-1, z)}{p(y, z)} + a_2 \frac{p(y, z-1)}{p(y, z)} + Ua_1^2 \frac{p(y-1, z-1)}{p(y, z)} \right],$$

où

$$\mathcal{F}_{t-1} = \sigma(X_{t-1}, X_{t-2}, \dots), U = (1 - a_1 - a_2)^{-1}$$

et

$$p(y, z) \equiv \mathbb{P}(X_{t-1} = y, X_{t-2} = z) = \exp[\lambda(a_1 - 2)U] \sum_{i=0}^{\min(y, z)} \frac{[\lambda U(1 - a_1)]^{y+z-2i} (\lambda U a_1)}{(y-i)! (z-i)! i!}.$$

1.2.3 Le modèle INAR(p)-DL

Du & Li (1991, [21]) ont présenté une nouvelle spécification du processus INAR(2), noté INAR(2)-DL. Ici, les auteurs supposent que les séries de comptage (ξ_{1k}) et (ξ_{2k}) sont indépendantes. Le processus INAR(2)-DL est stationnaire si $a_1 + a_2 < 1$.

De plus, ils montrent que

$$\mathbb{E}X_t = \frac{\lambda}{1 - a_1 - a_2}.$$

Cependant, la variance $V(X_t)$ n'est pas en général égale à $\lambda/(1 - a_1 - a_2)$.

Par conséquent, la distribution marginale de (X_t) n'est plus de Poisson si le bruit (ε_t) l'est.

Par contre, la structure de corrélation du processus INAR(p)-DL est identique à celle du processus AR(p) réel.

Alors, les estimateurs de Yule-Walker pour les paramètres a_1 et a_2 sont

$$\hat{a}_1 = \hat{\rho}(1) \left[\frac{1 - \hat{\rho}(2)}{1 - \hat{\rho}(1)^2} \right] \text{ et } \hat{a}_2 = \frac{\hat{\rho}(2) - \rho(1)^2}{1 - \hat{\rho}(1)^2}. \quad (1.22)$$

Ainsi, l'estimation de Yule-Walker de λ est

$$\hat{\lambda} = \bar{X}(1 - \hat{a}_1 - \hat{a}_2), \text{ où } \bar{X} \text{ est la moyenne empirique.} \quad (1.23)$$

De plus, la fonction de régression du modèle INAR-DL est linéaire

$$\mathbb{E}(X_t | \mathcal{F}_{t-1}) = a_1 X_{t-1} + a_2 X_{t-2} + \lambda. \quad (1.24)$$

Notons que le problème de l'estimation des paramètres d'un tel processus fut abordé indépendamment par Gauthier (1991, [29]).

1.2.4 Le modèle GINAR(p)

Dans le but d'enrichir la classe de modèles autorégressifs permettant l'analyse des séries chronologiques à valeurs entières, Gauthier & Latour (1994, [30]) Dion et al. (1995, [18]) et Latour (1998, [47]) ont considéré une version plus générale du modèle INAR(p), noté GINAR(p), basée sur l'opérateur généralisé de Steutel & Van Harn (1979, [59]).

Définition 1.2.2. (*L'opérateur d'amincissement généralisé, noté $a\star$*). Soit Z une variable aléatoire à valeurs entières positives ; ξ , une variable aléatoire à valeurs entières positives de moyenne finie a et de variance finie λ et (ξ_k) une suite de variables aléatoires indépendantes entre elles, indépendantes de Z et distribuées selon la même loi que ξ .

Alors, l'opérateur $a\star$ est défini par

$$a\star Z = \sum_{j=1}^Z \xi_j.$$

La suite (ξ_k) est aussi une série de comptage. De plus, supposons que $b\star$ est un autre opérateur d'amincissement généralisé de Z , basé sur une série de comptage (ψ_k) .

Les opérateurs $a\star$ et $b\star$ sont indépendants si, et seulement si, les séries de comptage (ξ_k) et (ψ_k) sont mutuellement indépendantes.

Définition 1.2.3. Soit $\{X_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires à valeurs entières positives ; $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, une suite de variables aléatoires i.i.d. à valeurs entières positives de moyenne finie μ_ε et de variance finie σ_ε^2 ; $p \in \mathbb{N}^*$ et $\{a_j\}_{j \in \{1,2,\dots,p\}}$, une suite de constantes telles que $\forall j, 0 \leq a_j < 1$ ($a_p > 0$) et $\sum_{j=1}^p a_j < 1$. Alors, $\{X_t\}_{t \in \mathbb{Z}}$ est un processus GINAR(p) si

$$X_t = \sum_{j=1}^p a_j \star X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}. \quad (1.25)$$

Ici, toutes les séries de comptage $(\xi_{jk})_{k \in \mathbb{N}^*}$ associées à $a_j\star$ pour $j = 1, 2, \dots, p$ sont indépendantes entre elles et indépendantes de ε_t . Elles sont de moyenne finie a_j et de variance finie λ_j . De plus, les auteurs (1994, [30]) n'exigent pas qu'elles soient de type Bernoulli, ce qui distingue nettement leurs résultats de ceux de Du & Li (1991, [21]). Le processus GINAR(p) possède une structure de corrélation identique à celle d'un processus AR(p) réel.

Sous l'hypothèse $\sum_{j=1}^p a_j < 1$, le processus GINAR(p) est stationnaire. Notons que Dion et al. (1995, [18]) ont également établi la stationnarité du GINAR(p). Ces derniers ont utilisé la théorie des processus de branchement multi-type.

Ainsi, les moments de premier et second ordre du processus GINAR(p) sont

$$\mathbb{E}X_t = \frac{\mu_\varepsilon}{1 - \sum_{j=1}^p a_j} = \mu_X. \quad (1.26)$$

et

$$V(X_t) = \mu_X \sum_{j=1}^p \lambda_j + \sum_{j=1}^p a_j \gamma(j) + \sigma_\varepsilon^2, \text{ où } \gamma(j) = \text{cov}(X_t, X_{t+j}), \forall t \in \mathbb{Z}. \quad (1.27)$$

Concernant l'estimation des paramètres de ce processus, les auteurs (1994, [30]) ont donné des résultats similaires aux résultats de Du& Li (1991, [21]).

De plus, Latour (1998, [47]) a montré que le corrélogramme, et par la suite la densité spectrale, lorsque celle-ci existe, du processus GINAR(p) (et par conséquent le processus INAR(p)-DL) coïncident avec ceux d'un processus AR(p) réel.

En effet, ce fait n'est pas clairement déclaré dans la littérature. À titre exemple, dans Du& Li (1991, [21]) il est écrit qu'un processus INAR(p) est similaire à un processus AR(p). De plus, Latour (1998, [47]) a donné quelques remarques concernant l'estimation des moindres carrés conditionnelle des paramètres. Notons que, les propriétés du modèle GINAR(p) ont été étudiées aussi par Da Silva et Oliveira (2005, 2006, [16], [17]) et Da Silva et Da Silva (2006, [15]). De plus, Drost et al. (2008, [20]) ont fourni un estimateur efficace des paramètres, et en particulier, ont montré que le modèle GINAR(p) (et par conséquent le processus INAR(p)-DL) possède la propriété de la normalité asymptotique locale (LAN).

1.2.5 Le modèle MGINAR(p)

Latour (1997, [46]) a introduit le modèle MGINAR(p), pour analyser les séries chronologiques à valeurs entières multivariées. Ce dernier est basé sur l'opérateur matriciel généralisé de Steutel & Van Harn.

Définition 1.2.4. (*L'opérateur matriciel d'amincissement généralisé, noté $A\star$*).

Soit $A\star = \{a_{ij}\star\}_{1 \leq i,j \leq d}$ une matrice $d \times d$ d'opérateurs d'amincissement généralisés. Soit Z un vecteur de variables aléatoires à valeurs entières positives de dimension d . L'effet de $A\star$ sur $Z = (Z_1, \dots, Z_d)^\tau$, noté $A\star Z$, est défini par

$$A\star \begin{pmatrix} Z_1 \\ \vdots \\ Z_d \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^d a_{1j}\star Z_j \\ \vdots \\ \sum_{j=1}^d a_{dj}\star Z_j \end{pmatrix}.$$

Ici, tous les opérateurs $\{a_{ij}\star\}_{1 \leq i,j \leq d}$ sont mutuellement indépendants.

Définition 1.2.5. Soit $\{X_t\}_{t \in \mathbb{Z}}$, une suite de vecteurs de variables aléatoires à valeurs entières positives de dimension d ; $p \in \mathbb{N}^*$ et $\{A_j \star\}_{j \in \{1,2,\dots,p\}}$, une suite d'opérateurs matriciels mutuellement indépendantes; $\{\varepsilon_t\}_{t \in \mathbb{Z}}$, une suite de vecteurs de dimension d de variables aléatoires i.i.d. à valeurs entières positives; de carré intégrable; indépendantes de tous les opérateurs. Alors $\{X_t\}_{t \in \mathbb{Z}}$ est un processus $MGINAR(p)$ si

$$X_t = \sum_{j=1}^p A_j \star X_{t-j} + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Latour (1997, [46]) a donné les conditions de stationnarité et de causalité d'un tel processus et a montré que la fonction d'autocovariance du $MGINAR(p)$ est identique à celle du processus autorégressif vectoriel standard à valeurs réelles, noté $VAR(p)$. Il en déduit que le processus $MGINAR(P)$ n'est autre qu'un processus $VAR(p)$. Par conséquent, la densité spectrale est directement trouvée et donne une bonne idée de la structure stochastique de $MGINAR(p)$. Par suite, l'auteur (1997, [46]) a considéré l'estimateur des moindres carrés conditionnels pour estimer les paramètres du modèle et a montré la consistance et la normalité asymptotique de l'estimateur.

1.3 Contributions de la thèse

Au cours de cette thèse, notre attention s'est portée uniquement sur les modèles autorégressifs à valeurs entières. Dans les paragraphes suivants, nous évoquons les limites des modèles, basés sur les opérateurs d'amincissement, mentionnés précédemment. Nous présentons une nouvelle classe de modèles autorégressifs à valeurs entières, basés sur l'opérateur d'arrondi à l'entier le plus près. Nous présentons les intérêts et les motivations d'introduire cette dernière. Finalement, nous fournissons un plan de notre travail.

1.3.1 Critiques des modèles basés sur les opérateurs d'amincissement

Les modèles autorégressifs à valeurs entières basés sur les opérateurs d'amincissement souffrent de plusieurs handicaps.

- Leur structure d'innovation est complexe, dépendant non seulement du bruit, mais également du choix des séries de comptage. Explicitement, nous considérons le modèle $INAR(2)$ défini par

$$X_t = a_1 \circ X_{t-1} + a_2 \circ X_{t-2} + \varepsilon_t, \quad \forall t \in \mathbb{Z},$$

où $0 \leq a_1 < 1$, $0 < a_2 < 1$ et $0 < a_1 + a_2 < 1$. Nous rappelons que

$$a_1 \circ X_{t-1} = \sum_{k=1}^{X_{t-1}} \xi_{1k} \text{ et } a_2 \circ X_{t-2} = \sum_{k=1}^{X_{t-2}} \xi_{2k},$$

où la série de comptage (ξ_{1k}) (resp. (ξ_{2k})) est une suite de variables aléatoires indépendantes, identiquement distribuées selon une loi de Bernoulli de paramètre a_1 (resp. a_2) et indépendante de X_{t-1} (resp. X_{t-2}). Ainsi, si (ξ_{1k}) et (ξ_{2k}) sont mutuellement indépendantes, alors le modèle INAR(2) possède la même structure de corrélation qu'un AR(2) réel (voir Section 1.2.3). Par contre, Si nous supposons une certaine dépendance entre les séries de comptage la structure de corrélation du processus INAR(2) devient identique à celle d'un ARMA(2,1) réel (voir Section 1.2.2).

- À cause de l'opérateur d'amincissement, ces modèles ne peuvent pas produire des autocorrélations négatives. Plus précisément, nous considérons le modèle INAR(1) défini par

$$X_t = a \circ X_{t-1} + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Rappelons que ce dernier possède la même structure de corrélation qu'un AR(1) réel. En particulier,

$$\rho(1) = a \in [0, 1[.$$

Ainsi, si nous disposons d'une série chronologique à valeurs entières avec $\hat{\rho}(1) < 0$, alors le modèle INAR(1) n'est approprié pour analyser cette série.

- En général, leur prévision à un pas basée sur l'espérance conditionnelle (la fonction de régression) est une valeur réelle. Ainsi, cette dernière n'appartient pas au support entier des observations de la série. À titre d'exemple, pour le processus GINAR(p) (par conséquent pour le processus INAR(p)-DL) la prévision, à l'instant T , basée sur la fonction de régression est donnée par

$$\hat{X}_T = \mathbb{E}[X_T | \mathcal{F}_{T-1}] = \sum_{j=1}^p a_j X_{T-j} + \lambda,$$

où $\mathcal{F}_{T-1} = \sigma\{X_{T-1}, \dots\}$ et $\mathbb{E}\varepsilon_t = \lambda$. \hat{X}_T n'est donc en pratique jamais entière.

1.3.2 Intérêts des modèles introduits

La plupart des modèles autorégressifs à valeurs entières existants dans la littérature ont pour but de modéliser des séries chronologiques à valeurs entières positives, en particulier les séries de comptage. Dans le cadre de cette thèse, notre intérêt s'est tourné vers la recherche

des modèles autorégressifs à valeurs entières qui peuvent produire aussi des observations négatives.

Pourquoi des modèles sur \mathbb{Z} ?

Une réponse naturelle et directe est que les séries chronologiques à valeurs entières ayant des observations négatives sont aussi fréquentes dans la pratique. À titre d'exemple, dans la Section 2.7 nous traitons la série des taux de variation annuels de la population Suédoise entre 1750 et 1849 où les observations varient entre -27 et 16 .

D'autre part, supposons que nous disposons d'une série chronologique à valeurs entières positives où nous détectons par exemple des fluctuations saisonnières. Pour éliminer la saisonnalité des observations (et ainsi rendre la série stationnaire) nous utilisons la technique de différenciation. Ainsi, nous sommes amenés à étudier une série chronologique à valeurs entières possédant des observations négatives.

Par ailleurs, supposons que nous disposons d'une série chronologique à valeurs entières positives assez élevées, nous souhaitons réduire les variations de ces observations tout en en gardant leur nature entière. En d'autres termes, en utilisant des techniques similaires au centrage dans le cas des observations réelles, nous aboutissons à des observations entières à faibles variations qui oscillent autour de 0. Encore une fois, nous sommes amenés à traiter une série chronologique à valeurs entières possédant des observations négatives.

Le modèle **RINAR**

Pour générer des observations à valeurs entières, nous envisageons l'opérateur d'arrondi à l'entier le plus près. Ce dernier est souvent utilisé pour récolter des observations entières à partir des données réelles (par exemple, les données météorologique).

Ainsi, nous introduisons le modèle suivant

$$X_t = \left\langle \sum_{j=1}^p \alpha_j X_{t-j} + \lambda \right\rangle + \varepsilon_t, \quad (1.28)$$

où $\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près, (ε_t) est une suite de variables aléatoires i.i.d. centrées à valeurs dans \mathbb{Z} , λ et les α_j sont des paramètres réels.

Ce modèle est appelé **RINAR(p)** (pour Rounded INteger-valued AutoRegressif).

Notons que l'opérateur d'arrondi peut être interprété comme une fonction de censure sur le modèle $\text{AR}(p)$ réel et λ comme la moyenne du bruit non-centré ($\varepsilon'_t = \varepsilon_t + \lambda$).

Les avantages du modèle introduit

Comparé aux modèles autorégressifs à valeurs entières basés sur les opérateurs d'arrondissement le modèle RINAR présente quelques avantages.

- Sa structure d'innovation est simple, générée uniquement par le bruit (ε_t) .
- Par construction RINAR(p) peut analyser des séries chronologiques avec des valeurs négatives, une situation qui n'est couverte par aucun modèle INAR.
- Sa prévision à un pas, basée sur l'espérance conditionnelle, est donnée par

$$\hat{X}_T = \mathbb{E}(X_T \mid X_s, s \leq T-1) = \left\langle \sum_{j=1}^p \alpha_j X_{T-j} + \lambda \right\rangle.$$

Cette dernière est une valeur entière par construction du modèle.

- Nous verrons également que le modèle RINAR(p) peut produire des autocorrélations aussi riches que celles d'un AR(p) réel, y compris les autocorrélations négatives.

1.3.3 Plan de la thèse

L'introduction de nouvelles classes de modèles autorégressifs à valeurs entières présente l'objet de nos recherches au cours de cette thèse. Dans ce but, nous nous sommes basés sur l'opérateur d'arrondi à l'entier le plus près qui nous semble assez intuitif et naturel pour générer des observations à valeurs entières. Les modèles proposés dans cette thèse exploitent un domaine rarement examiné : la modélisation des séries chronologiques à valeurs entières ayant des observations négatives. Par ailleurs, comparés aux modèles existants dans la littérature ces nouveaux modèles possèdent plusieurs avantages. Ainsi, la thèse est organisée de la manière suivante.

L'étude théorique du modèle arrondi autorégressif à valeurs entières du premier ordre, RINAR(1), fait l'objet du **Chapitre 2** de cette thèse. Ainsi, après une présentation des propriétés de l'opérateur d'arrondi à l'entier le plus près, nous introduisons le modèle RINAR(1). Ensuite, nous donnons les conditions de stationnarité et d'ergodicité de ce processus. Spécifiquement, nous signalons que comme pour un processus AR(1) réel, la valeur absolue du coefficient de régression doit être inférieure strictement à 1. Par ailleurs, nous présentons quelques propriétés générales du modèle introduit. En particulier, nous nous intéressons aux calculs de la moyenne et des coefficients d'autocorrélation du processus RINAR(1) stationnaire. Nous mentionnons qu'à cause de l'opérateur d'arrondi un calcul explicite de ces paramètres n'est pas possible. Cependant, nous arrivons à encadrer ces

paramètres en fonction de leurs valeurs homologues dans le cas d'un processus AR(1) réel. D'autre part, nous examinons le problème d'identifiabilité, dû aussi à l'opérateur d'arrondi, et la consistance forte de l'estimateur des moindres carrés proposé pour l'estimation des paramètres du modèle. Dans la suite, nous proposons une méthode numérique pour calculer ce dernier. Finalement, nous analysons deux exemples de série chronologique à valeurs entières bien connus, avec le processus RINAR(1).

Le **Chapitre 3** est consacré à l'étude du modèle arrondi autorégressif à valeurs entières d'ordre supérieur, noté RINAR(p). Ce dernier est une extension naturelle et directe du modèle RINAR(1) étudié dans le Chapitre 2. En premier lieu, nous imposons des conditions pour assurer la stationnarité et l'ergodicité du processus. Ensuite, nous proposons l'estimateur des moindres carrés pour estimer les paramètres du modèle RINAR(p). Puis, nous examinons le problème d'identifiabilité, dû à l'opérateur d'arrondi, et la consistance de ce dernier. Par ailleurs, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Cette dernière n'est autre qu'une généralisation de la méthode numérique considérée dans le Chapitre 2. Pour en finir, nous présentons une application réelle traitée avec un modèle RINAR(p).

Dans le **Chapitre 4** nous étudions le modèle RINVAR dédié à l'analyse des séries chronologiques à valeurs entières multivariées. En particulier, nous nous concentrons sur le modèle du premier ordre, noté RINVAR(1). Premièrement, nous donnons les conditions de stationnarité et d'ergodicité du processus. Ensuite, nous examinons le problème d'identifiabilité, dû à l'opérateur d'arrondi, et la consistance de l'estimateur des moindres carrés proposé pour estimer les paramètres du modèle. D'autre part, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Nous terminons le chapitre par une application réelle traitée avec un modèle RINVAR(1).

L'étude du modèle arrondi centré autorégressif à valeurs entières du premier ordre fait l'objet du **Chapitre 5**. Ce dernier est appelé RINAR(1) centré et permet d'analyser des séries chronologiques à valeurs entières dont les observations oscillent autour de 0 (et ainsi de moyenne nulle). La stationnarité et l'ergodicité de ce modèle sont assurées dans les mêmes conditions que celles pour le modèle RINAR(1). Le RINAR(1) centré ne possède pas un problème d'identifiabilité. Ainsi, l'estimateur des moindres carrés considéré pour l'estimation du paramètre du modèle est fortement consistant. Nous introduisons des hypothèses supplémentaires qui assurent que la moyenne du modèle RINAR(1) centré est nulle. Notons que ce modèle souffre des mêmes problèmes techniques que pour le modèle

RINAR(1). En particulier, nous n'arrivons pas à calculer explicitement ces coefficients d'autocorrélation mais nous pouvons les encadrer. Par ailleurs, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Suite à une étude de simulation qui compare les coefficients d'autocorrélation du modèle RINAR(1) centré et ceux d'un AR(1) réel, nous distinguons deux comportements différents dépendant de la nature du bruit considéré (grand ou petit). Finalement, nous comparons le modèle RINAR(1) centré étudié dans ce chapitre et le modèle RINAR(1) recentré.

Dans le **Chapitre 6**, nous introduisons un processus autorégressif du premier ordre basé sur l'opérateur d'arrondi pour analyser des séries chronologiques à valeurs entières positives, noté PRINAR(1). En premier lieu, nous introduisons les conditions qui assurent la positivité, la stationnarité et l'ergodicité de ce modèle. Par la suite, nous donnons le résultat principal concernant la consistance forte de l'estimateur des moindres carrés considéré pour estimer les paramètres du modèle. Puis, nous étudions les propriétés générales du modèle. En particulier, nous trouvons des encadrements pour sa moyenne et ses coefficients d'autocorrélation. Suite à une étude de simulation concernant ces derniers, nous remarquons beaucoup de détails curieux dans le cas où bruit utilisé est petit. Finalement, nous donnons les conditions nécessaires pour que la loi stationnaire du processus PRINAR(1) possède un support fini.

Chapitre 2

Le modèle RINAR(1)

Sommaire

2.1	Quelques définitions	20
2.2	Stationnarité et ergodicité du processus RINAR(1)	24
2.3	Propriétés du processus RINAR(1)	26
2.4	Estimation des paramètres	29
2.4.1	La consistance forte de l'estimateur des moindres carrés quand le coefficient de régression α_0 est irrationnel	34
2.4.2	La consistance forte de l'estimateur des moindres carrés quand le coefficient de régression α_0 est rationnel	43
2.5	Méthode numérique pour calculer $\hat{\theta}_n$	57
2.5.1	Initialisation	58
2.5.2	Recherche dichotomique successive	58
2.5.3	Étude de simulation	61
2.6	Analyse des données d'O'Donovan	65
2.6.1	Le modèle AR(1) d'O'Donovan	67
2.6.2	Ajustement à un modèle RINAR(1)	69
2.6.3	Comparaison et commentaires	71
2.7	Analyse des taux de variation annuels de la population Suédoise	71
2.7.1	Ajustement à un modèle AR(1)	74
2.7.2	Ajustement à un modèle RINAR(1)	74
2.7.3	L'approche de McCleary & Hay	76
2.7.4	Comparaison et commentaires	77

Pour générer des observations à valeurs entières nous considérons l'opérateur d'arrondi à l'entier le plus près. Ce chapitre est consacré à l'étude théorique du modèle RINAR(1) (pour rounded integer-valued autoregressive), basé sur l'opérateur d'arrondi.

Dans le premier paragraphe nous présentons quelques propriétés utiles de l'opérateur d'arrondi à l'entier le plus près. Dans le deuxième paragraphe, nous introduisons le modèle RINAR(1), nous donnons les conditions de stationnarité et d'ergodicité du processus. Par suite, dans le troisième paragraphe, nous présentons quelques propriétés générales du modèle. En particulier, nous nous intéressons aux calculs des coefficients d'autocorrélation du processus RINAR(1) stationnaire. Dans le quatrième paragraphe, nous examinons le problème d'identifiabilité, dû à l'opérateur d'arrondi, et la consistance forte de l'estimateur des moindres carrés proposé pour l'estimation des paramètres du modèle. Dans le cinquième paragraphe, nous proposons une méthode numérique pour calculer ce dernier. Finalement, dans les deux derniers paragraphes, nous analysons deux exemples de série chronologique à valeurs entières bien connus, avec le processus RINAR(1).

2.1 Quelques définitions

Nous introduisons dans ce paragraphe diverses notations, définitions et propriétés qui sont utilisées dans la suite de ce chapitre. Tout d'abord, rappelons les notations classiques suivantes

$$\mathbb{N} = \{x \in \mathbb{Z} : x \geq 0\}, \mathbb{N}^* = \{x \in \mathbb{Z} : x > 0\} \text{ et } \mathbb{Z}^- = \{x \in \mathbb{Z} : x \leq 0\}.$$

$\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près. Soit a un nombre réel.

Notons que la valeur $\langle a \rangle$ est clairement définie partout, sauf si $a = k + \frac{1}{2}$, où $k \in \mathbb{Z}$.

Par convention, nous prenons pour tout $k \in \mathbb{N}$ et $p \in \mathbb{Z}^-$

$$\langle k + \frac{1}{2} \rangle = k + 1 \text{ et } \langle p - \frac{1}{2} \rangle = p - 1 \text{ (voir Figure 2.1)}. \quad (2.1)$$

De plus la fonction arrondi, définie sur \mathbb{R} et à valeurs dans \mathbb{Z} , est impaire.

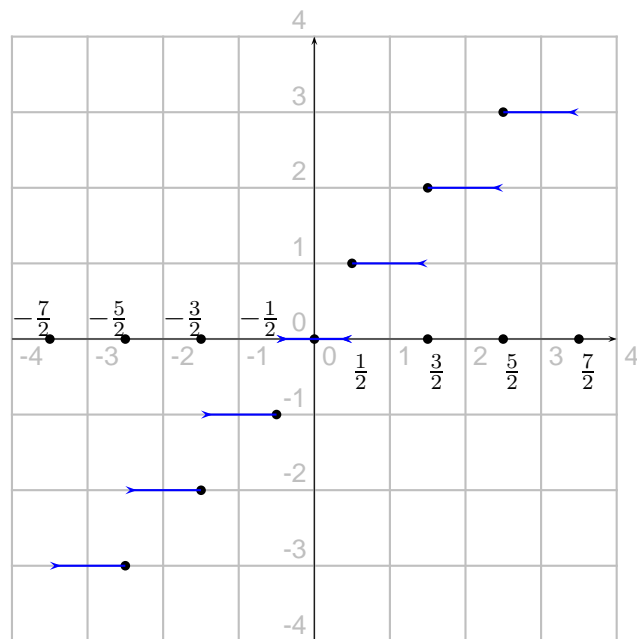


FIG. 2.1 – La fonction arrondi

Soit $\{a\}$ la *partie fractionnaire* du nombre réel a , nous avons donc $\{a\} \in [0, 1[$.

Ici, nous considérons que la partie fractionnaire d'un réel négatif est égale à celle de sa valeur absolue (voir Figure 2.2), c.à.d.

$$\forall a \in \mathbb{R}, \{a\} = \{-a\} = \{|a|\}, \text{ par exemple } \{1.23\} = \{-1.23\} = 0.23. \quad (2.2)$$

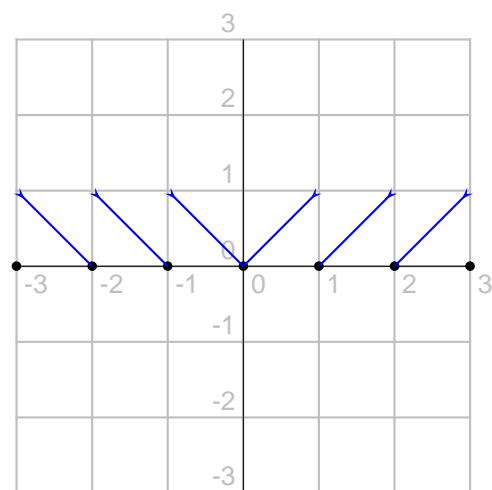


FIG. 2.2 – La fonction partie fractionnaire

La fonction signe, notée s , est définie par

$$s(a) = \begin{cases} 1, & \text{si } a \geq 0, \\ -1, & \text{si } a < 0. \end{cases} \quad (2.3)$$

Soit $[a]$ la partie entière de $a \in \mathbb{R}$, par exemple $[2.8] = 2$ et $[-1.8] = -1$ (voir Figure 2.3).

Il en résulte, pour tout $a \in \mathbb{R}$,

$$a = [a] + s(a) \{a\}. \quad (2.4)$$

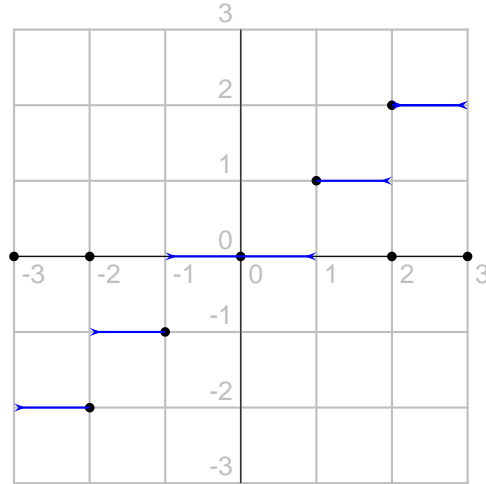


FIG. 2.3 – La fonction partie entière

Ainsi, pour tout $a \in \mathbb{R}$, nous trouvons

$$a = \langle a \rangle + s(a) \left(\{a\} - \mathbb{1}_{\{a\} \geq \frac{1}{2}} \right). \quad (2.5)$$

Le lemme suivant rapporte quelques propriétés générales. Ce dernier est une conséquence directe des équations (2.4) et (2.5).

Lemme 2.1.1. *Soient $x \in \mathbb{R}$, $k \in \mathbb{Z}$ et $a, b \geq 0$.*

$$1. \quad |\langle x \rangle - [x]| = \mathbb{1}_{\{x\} \geq \frac{1}{2}}.$$

$$2. \quad |\langle x \rangle - x| \leq \frac{1}{2}.$$

$$3. \quad |\langle x \rangle| = \langle |x| \rangle \leq |x| + \frac{1}{2}.$$

$$4. \quad \langle x + k \rangle = k + \langle x \rangle.$$

$$5. \quad \langle a + b \rangle = c + \langle \{a\} + \{b\} \rangle, \text{ où } c = \langle a \rangle + \langle b \rangle - \mathbb{1}_{\{a\} \geq \frac{1}{2}} - \mathbb{1}_{\{b\} \geq \frac{1}{2}}.$$

6.

$$\{a + b\} = \{\{a\} + \{b\}\} = \begin{cases} \{a\} + \{b\}, & \text{si } \{a\} + \{b\} < 1, \\ \{a\} + \{b\} - 1, & \text{si } \{a\} + \{b\} \geq 1. \end{cases}$$

$$7. \quad \langle a \rangle = [a] + \langle \{a\} \rangle.$$

8.

$$\langle a + b \rangle = [a + b] + \langle \{\{a\} + \{b\}\} \rangle = \begin{cases} [a + b], & \text{si } 0 \leq \{a\} + \{b\} < \frac{1}{2}, \\ & \text{ou } 1 \leq \{a\} + \{b\} < \frac{3}{2}, \\ [a + b] + 1, & \text{si } \frac{1}{2} \leq \{a\} + \{b\} < 1, \\ & \text{ou } \frac{3}{2} \leq \{a\} + \{b\} < 2. \end{cases}$$

9. Si $[a] = [b]$, alors

$$\langle a \rangle - \langle b \rangle = \langle \{a\} \rangle - \langle \{b\} \rangle = \begin{cases} 0, & \text{si } \{a\} < \frac{1}{2} \text{ et } \{b\} < \frac{1}{2}, \\ & \text{ou } \{a\} \geq \frac{1}{2} \text{ et } \{b\} \geq \frac{1}{2}, \\ 1, & \text{si } \{a\} \geq \frac{1}{2} \text{ et } \{b\} < \frac{1}{2}, \\ -1, & \text{si } \{a\} < \frac{1}{2} \text{ et } \{b\} \geq \frac{1}{2}. \end{cases}$$

10. Supposons $\langle a \rangle = \langle b \rangle$ et $\{a\} \in \left[0, \frac{1}{2}\right[$. Nous distinguons deux cas :

- Si $[a] = [b]$, alors $\{b\} \in \left[0, \frac{1}{2}\right[$.

- Si $[a] = [b] + 1$, alors $\{b\} \in \left[\frac{1}{2}, 1\right[$.

11. Supposons que $\langle a \rangle = \langle b \rangle$ et $\{a\} \in \left[\frac{1}{2}, 1\right[$. Nous distinguons deux cas :

- Si $[a] = [b]$, alors $\{b\} \in \left[\frac{1}{2}, 1\right[$.

- Si $[a] = [b] - 1$, alors $\{b\} \in \left[0, \frac{1}{2}\right[$.

2.2 Stationnarité et ergodicité du processus RINAR(1)

Le processus arrondi à valeurs entières autorégressif de premier ordre, noté RINAR(1), a été introduit par Kachour & Yao (2009, [43]) pour analyser des séries chronologiques à valeurs entières.

Définition 2.2.1. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Le processus $\{X_t, t \in \mathbb{Z}\}$ est dit un processus RINAR(1) si pour tout t , il possède la présentation suivante*

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t, \quad (2.6)$$

où $\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près, (ε_t) est une suite de variables aléatoires i.i.d. centrées à valeurs dans \mathbb{Z} , définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, λ et α sont des paramètres réels.

Notons que l'opérateur d'arrondi peut être interprété comme une fonction de censure sur le modèle AR(1) et λ comme la moyenne du bruit non-centré ($\varepsilon'_t = \varepsilon_t + \lambda$).

Comme le bruit (ε_t) est une suite de variables aléatoires i.i.d., le processus (X_t) défini par (2.6), forme une chaîne de Markov homogène avec un espace d'états $E = \mathbb{Z}$ et une probabilité de transition

$$\pi(x, y) = \mathbb{P}\{\varepsilon_1 = y - f(x, \theta)\}, \quad x, y \in E, \quad (2.7)$$

où la fonction de régression $f(x, \theta) = \langle \alpha x + \lambda \rangle$, pour tout $x \in E$ et $\theta = (\alpha, \lambda) \in \Theta$, l'espace des paramètres.

Proposition 2.2.1. *Soit $\theta = (\alpha, \lambda) \in \Theta$ fixé. Supposons que :*

1. *La chaîne de Markov (X_t) est irréductible ;*
2. *pour un certain $k > 1$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;*
3. *$|\alpha| < 1$.*

Alors,

1. *(X_t) possède une unique mesure de probabilité invariante, notée μ . De plus, μ possède un moment d'ordre k .*
2. *Pour tout $x \in E$ et $f \in L^1(\mu)$ nous avons*

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \longrightarrow \mu(f), \quad \mathbb{P}_x \text{ p.s.}$$

où \mathbb{P}_x représente la probabilité conditionnelle $\mathbb{P}(\cdot \mid X_0 = x)$.

Preuve. Pour tout $x_0 \in E = \mathbb{Z}$ nous définissons la mesure empirique par :

$$\mu_n(\cdot) = \frac{1}{n} [\pi^1(x_0, \cdot) + \cdots + \pi^n(x_0, \cdot)],$$

où π^n représente la probabilité de transition de la chaîne de Markov en n pas.

L'idée de la preuve est de montrer que cette suite de mesure de probabilité (μ_n) possède une sous-suite qui converge vers une certaine mesure de probabilité μ .

Par construction, cette limite sera la mesure de probabilité stationnaire de la chaîne de Markov (X_t) . Nous définissons sur $E = \mathbb{Z}$ la fonction V par :

$$x \mapsto V(x) = |x|^k, \quad \text{où } k > 1.$$

Comme V est positive et $\lim_{|x| \rightarrow \infty} V(x) = \infty$, V est donc une fonction de Lyapunov.

Ainsi, d'après l'inégalité de Minkowski, nous obtenons

$$\begin{aligned} \pi V(x)^{\frac{1}{k}} &= \left(\mathbb{E}_x |\langle \alpha X_0 + \lambda \rangle + \varepsilon_1|^k \right)^{\frac{1}{k}} \\ &\leq \mathbb{E}_x \left(|f(X_0; \theta)|^k \right)^{\frac{1}{k}} + c \\ &= \mathbb{E}_x \left(|\langle \alpha X_0 + \lambda \rangle|^k \right)^{\frac{1}{k}} + c. \end{aligned}$$

où $c = (\mathbb{E} |\varepsilon_1|^k)^{\frac{1}{k}} < \infty$, d'après l'hypothèse 2. De plus, nous avons

$$|f(X_0; \theta)| = |\langle \alpha X_0 + \lambda \rangle| \leq |\alpha X_0| + |\lambda| + \frac{1}{2}.$$

Il en résulte

$$\pi V(x)^{\frac{1}{k}} \leq |\alpha| |x| + b = |x| \left(|\alpha| + \frac{b}{|x|} \right),$$

où $b = c + |\lambda| + \frac{1}{2}$. Par conséquent,

$$\frac{\pi V(x)}{V(x)} \leq \left(|\alpha| + \frac{b}{|x|} \right)^k.$$

Comme $|\alpha| < 1$, nous obtenons

$$\limsup_{|x| \rightarrow \infty} \frac{\pi V(x)}{V(x)} \leq |\alpha|^k < 1. \quad (2.8)$$

Par suite, nous obtenons l'équivalence suivante

$$\limsup_{|x| \rightarrow \infty} \frac{\pi V(x)}{V(x)} \leq a_0 = |\alpha|^k < 1 \Leftrightarrow \pi V(x) \leq \gamma V(x) + \delta, \quad \text{où } 0 \leq \gamma < 1 \text{ et } \delta < \infty. \quad (2.9)$$

Il en découle

$$\limsup_n \mu_n V \leq l < \infty.$$

Le critère de Lyapunov est satisfait. Ainsi, la suite (μ_n) est tendue et pour tout point limite μ nous avons $\mu V \leq l < \infty$ (voir Duflo [22], 1997, Prop. 2.1.6, p. 41).

Par suite, μ est une mesure π -invariante. De plus d'après l'hypothèse 1, (X_t) est une chaîne de Markov récurrente positive et par conséquent μ est unique. La conclusion 2 de la proposition est une conséquence directe du théorème ergodique classique pour les chaînes de Markov. \square

Les hypothèses de la Proposition 2.2.1 sont toutes naturelles et aisément vérifiées. Par exemple, si la loi du bruit ε_t charge tous les points de $E = \mathbb{Z}$, c.à.d. $\forall k \in \mathbb{Z}, \mathbb{P}\{\varepsilon_1 = k\} > 0$, l'hypothèse 1 concernant l'irréductibilité de la chaîne est assurée. Notons que l'hypothèse 3 est équivalente à la condition qui assure la stationnarité du modèle AR(1) réel.

D'autre part, notons que le support de la mesure de probabilité invariante μ est l'espace des états $E = \mathbb{Z}$ de la chaîne (X_t) . En d'autres termes, pour tout $x \in \mathbb{Z}$, nous avons $\mu(x) > 0$.

2.3 Propriétés du processus RINAR(1)

Dans ce paragraphe nous supposons que les hypothèses de la proposition 2.2.1 sont satisfaites. Par conséquent le processus RINAR(1) défini par (2.6) est stationnaire.

Rappelons que μ représente la mesure invariante de ce dernier.

Ainsi, nous notons

$$m = \mathbb{E}X_t, \forall t \in \mathbb{Z} \text{ et } \rho(k) = \frac{\text{cov}(X_t, X_{t+k})}{V(X_t)}, \forall t \in \mathbb{Z} \text{ et } k \in \mathbb{N}.$$

À cause de l'opérateur d'arrondi l'étude de la loi stationnaire et du corrélogramme du modèle RINAR(1) est compliquée. À titre d'exemple, nous n'arrivons pas à calculer explicitement la moyenne m ou le coefficient d'autocorrélation du premier ordre $\rho(1)$ du processus.

Nous donnons ci-dessous quelques encadrements pour ces paramètres du processus.

Proposition 2.3.1. *Sous les hypothèses de la Proposition 2.2.1, nous avons*

$$\frac{\lambda - \frac{1}{2}}{1 - \alpha} \leq m \leq \frac{\lambda + \frac{1}{2}}{1 - \alpha}. \quad (2.10)$$

Preuve. Nous définissons le processus suivant

$$W_t = s(\alpha X_t + \lambda) \left(\{\alpha X_t + \lambda\} - \mathbb{1}_{\{\alpha X_t + \lambda\} \geq \frac{1}{2}} \right), \forall t \in \mathbb{Z}. \quad (2.11)$$

Ainsi, RINAR(1) peut être écrit sous la forme suivante

$$X_t = \alpha X_{t-1} + \lambda - W_{t-1} + \varepsilon_t. \quad (2.12)$$

Il est simple de vérifier que $|W_t| \leq \frac{1}{2}$. De plus, comme (X_t) est stationnaire, le processus (W_t) l'est aussi. Par conséquent,

$$(1 - \alpha)\mathbb{E}X_t = \lambda - \mathbb{E}W_t,$$

par suite

$$|\mathbb{E}X_t - \frac{\lambda}{1 - \alpha}| = \frac{|\mathbb{E}W_t|}{1 - \alpha} \leq \frac{\frac{1}{2}}{1 - \alpha}.$$

Il en résulte,

$$\frac{\lambda - \frac{1}{2}}{1 - \alpha} \leq m \leq \frac{\lambda + \frac{1}{2}}{1 - \alpha}.$$

□

Donc, d'après (2.10), nous avons

$$m(1 - \alpha) - \frac{1}{2} \leq \lambda \leq m(1 - \alpha) + \frac{1}{2} \quad (2.13)$$

Notons que si $\lambda \in \mathbb{Z}$, nous obtenons

$$\lambda = \langle m(1 - \alpha) \rangle. \quad (2.14)$$

Nous remarquons que $|\lambda| \leq \frac{1}{2}$ est une condition nécessaire mais pas suffisante pour que $m = 0$. Donc si $|\lambda| \leq \frac{1}{2}$, alors l'inéquation (2.10) devient

$$|m| \leq \frac{1}{1 - \alpha}.$$

Donc,

$$\begin{cases} \text{si } \alpha \rightarrow 0, & \text{alors } |m| \leq 1, \\ \text{si } \alpha \rightarrow -1, & \text{alors } |m| \leq \frac{1}{2}, \\ \text{si } \alpha \rightarrow 1, & \text{alors } \frac{1}{1 - \alpha} \rightarrow \infty. \end{cases}$$

Proposition 2.3.2. *Supposons que les hypothèses de la Proposition 2.2.1 soient vérifiées.*

Si $m \in \mathbb{Z}$, alors pour tout $j \in \mathbb{N}^$*

$$|\rho(j) - \alpha^j| \leq \frac{1}{2} \sum_{i=0}^{j-1} |\alpha|^i.$$

Preuve. Nous considérons le processus centré $Y_t = X_t - m$ ($\mathbb{E}Y_t = 0$). Comme $m \in \mathbb{Z}$, le processus (Y_t) est à valeurs dans \mathbb{Z} . Ce dernier peut être écrit sous la forme suivante

$$X_t - m = \langle \alpha(X_{t-1} - m) + \alpha m + \lambda \rangle - m + \varepsilon_t, \quad (2.15)$$

et par suite

$$Y_t = \langle \alpha Y_{t-1} + \alpha m + \lambda \rangle - m + \varepsilon_t. \quad (2.16)$$

Finalement, comme $m \in \mathbb{Z}$, nous obtenons

$$Y_t = \langle \alpha Y_{t-1} + c \rangle + \varepsilon_t, \quad (2.17)$$

où $|c| = |\lambda - m(1 - \alpha)| \leq \frac{1}{2}$. Rappelons que ε_t est indépendante de $\mathcal{F}_t = \sigma\{X_t, X_{t-1}, \dots\}$.

Ainsi, comme $\forall t \in \mathbb{Z}, \mathbb{E}Y_t = 0$ et $|Y_t| \leq Y_t^2$, nous obtenons

$$\text{cov}(Y_t, Y_{t+1}) = \alpha V(Y_t) + \mathbb{E}Y_t(\langle \alpha Y_t + c \rangle - (\alpha Y_t + c)). \quad (2.18)$$

Notons que $|\langle x \rangle - x| \leq \frac{1}{2}$, pour tout $x \in \mathbb{R}$. Par conséquent,

$$|\rho(1) - \alpha| = \frac{|\mathbb{E}Y_t(\langle \alpha Y_t + c \rangle - (\alpha Y_t + c))|}{V(Y_t)} \leq \frac{1}{2} \frac{\mathbb{E}|Y_t|}{\mathbb{E}Y_t^2} \leq \frac{1}{2}. \quad (2.19)$$

Maintenant, supposons que $j = 2$. Nous avons,

$$\begin{aligned} \text{cov}(Y_t, Y_{t+2}) &= \text{cov}(Y_t, \langle \alpha Y_{t+1} + c \rangle + \varepsilon_{t+2}) \\ &= \text{cov}(Y_t, \langle \alpha Y_{t+1} + c \rangle) \\ &= \text{cov}(Y_t, \alpha Y_{t+1} + c + \langle \alpha Y_{t+1} + c \rangle - (\alpha Y_{t+1} + c)) \\ &= \alpha \text{cov}(Y_t, Y_{t+1}) + \mathbb{E}Y_t(\langle \alpha Y_{t+1} + c \rangle - (\alpha Y_{t+1} + c)). \end{aligned}$$

Ainsi, d'après (2.18), nous obtenons

$$\begin{aligned} \text{cov}(Y_t, Y_{t+2}) &= \alpha^2 V(Y_t) + \alpha \mathbb{E}Y_t(\langle \alpha Y_t + c \rangle - (\alpha Y_t + c)) \\ &\quad + \mathbb{E}Y_t(\langle \alpha Y_{t+1} + c \rangle - (\alpha Y_{t+1} + c)). \end{aligned}$$

Donc,

$$\begin{aligned} |\rho(2) - \alpha^2| &\leq \frac{1}{2} \frac{\mathbb{E}|Y_t|}{\mathbb{E}Y_t^2} (1 + |\alpha|) \\ &\leq \frac{1}{2} (1 + |\alpha|). \end{aligned}$$

□

Proposition 2.3.3. *Supposons que les hypothèses de la Proposition 2.2.1 soient vérifiées.*

Si $m \notin \mathbb{Z}$, alors pour tout $j \in \mathbb{N}^$*

$$|\rho(j) - \alpha^j| \leq \frac{1}{2} \left(1 + \frac{\mu([m])}{\Sigma'} \right) \sum_{i=0}^{j-1} |\alpha|^i,$$

où μ est la loi stationnaire du processus (X_t) , $[\cdot]$ représente l'opérateur partie entière, et $\Sigma' = \sum_{i \neq [m]} \mu(i) |i - m|$.

Preuve. Ici, le processus (Y_t) , défini par (2.16), est à valeurs dans \mathbb{R} .

Par suite $\forall t \in \mathbb{Z}$, $|Y_t| \not\leq Y_t^2$. Pour tout $j \in \mathbb{N}^*$, d'après la preuve de la Proposition 2.3.2, nous avons

$$|\rho(j) - \alpha| \leq \frac{1}{2} \frac{\mathbb{E} |X_t - m|}{\mathbb{E} (X_t - m)^2} \sum_{i=0}^{j-1} |\alpha|^i.$$

Il est clair que si $X_t = [m]$, alors $|Y_t| = |[m] - m| < 1$ et par suite $|Y_t| > Y_t^2$.

Nous avons,

$$\mathbb{E} |X_t - m| = \mu([m]) |[m] - m| + \Sigma',$$

où $\Sigma' = \sum_{i \neq [m]} \mu(i) |i - m|$. Ainsi, nous obtenons

$$\mathbb{E} |X_t - m| \leq \Sigma' \left(1 + \frac{\mu([m])}{\Sigma'} \right)$$

D'autre part, nous avons

$$\mathbb{E} (X_t - m)^2 = \mu([m]) ([m] - m)^2 + \Sigma.$$

où $\Sigma = \sum_{i \neq [m]} \mu(i) (i - m)^2$. De plus, nous avons $\Sigma' \leq \Sigma \leq \mathbb{E} (X_t - m)^2$.

Donc,

$$\frac{\mathbb{E} |X_t - m|}{\mathbb{E} (X_t - m)^2} \leq 1 + \frac{\mu([m])}{\Sigma'}.$$

□

2.4 Estimation des paramètres

Dans ce paragraphe nous considérons l'estimateur des moindres carrés pour estimer les paramètres du processus RINAR(1). À cause de l'opérateur d'arrondi l'identifiabilité du modèle n'est pas standard. Nous distinguons deux cas dépendants de la nature de la vraie valeur du paramètre α . Par suite, nous examinons la consistance forte de l'estimateur des moindres carrés pour les deux cas.

Nous considérons le modèle RINAR(1) défini par

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t = f(X_{t-1}; \theta) + \varepsilon_t, \quad (2.20)$$

avec $f(x; \theta) = \langle \alpha x + \lambda \rangle$ et $\theta = (\alpha, \lambda) \in \Theta$, compact sous-ensemble de $] -1, 1[\times \mathbb{R}$.

Quelques notations sont nécessaires. Soient $\theta_0 = (\alpha_0, \lambda_0)$ la vraie valeur du paramètre θ et \mathbb{P}_{θ_0} la distribution de la probabilité de la chaîne (X_t) sous le vrai modèle.

De plus, toute convergence $\xrightarrow{p.s.}$ signifie une convergence *p.s.* sous $\mathbb{P}_{\theta_0, x}$, ce qui se tient indépendamment de l'état initial x .

Soient X_0, X_1, \dots, X_n des observations du processus RINAR(1). Pour l'estimation du paramètre θ , nous considérons l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \varphi_n(\theta), \quad (2.21)$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n [X_t - f(X_{t-1}; \theta)]^2. \quad (2.22)$$

Soit l'hypothèse suivante.

Hypothèse [H1]

1. Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible,
2. Pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$,
3. α_0 , la vraie valeur de α , appartient à l'intervalle $] -1, 1[$,
4. l'espace des paramètres Θ , est un compact sous-ensemble de $] -1, 1[\times \mathbb{R}$.

Supposons que [H1] soit vérifiée. Donc, sous \mathbb{P}_{θ_0} et d'après la Proposition 2.2.1, (X_t) possède une unique mesure invariante μ_{θ_0} telle que $\mu_{\theta_0}(|\cdot|^k) < \infty$ avec $k \geq 2$. Il en résulte que la chaîne double (Y_t) avec $Y_t = (X_{t-1}, X_t)$ possède des propriétés similaires.

Son noyau de transition Π_{θ_0} est égale

$$\Pi_{\theta_0}((x, z), (x', z')) = \pi_{\theta_0}(z, z') \mathbb{1}_{z=x'}, \quad (x, z), (x', z') \in E^2,$$

où π_{θ_0} est la probabilité de transition de la chaîne (X_t) sous le vrai modèle.

La chaîne (Y_t) a aussi une unique mesure invariante $\sigma_{\theta_0} = \mu_{\theta_0} \otimes \pi_{\theta_0}$, c.à.d.

$$\sigma_{\theta_0}((x, z)) = \mu_{\theta_0}(x) \pi_{\theta_0}(x, z), \quad (x, z) \in E^2. \quad (2.23)$$

Comme $\mu_{\theta_0}(|\cdot|^k) < \infty$, il s'ensuit $\sigma_{\theta_0}(\|\cdot\|^k) < \infty$ pour toute norme $\|\cdot\|$ sur E^2 .

Soient les fonctions :

$$g_{\theta}(y) = [z - f(x; \theta)]^2, \quad \forall y = (x, z) \in E^2 \text{ et } \theta \in \Theta. \quad (2.24)$$

$$K(\theta) = \sigma_{\theta_0} g_{\theta}(\cdot), \quad \forall \theta \in \Theta. \quad (2.25)$$

Soit \mathbb{P}_n la mesure empirique générée par les observations Y_1, \dots, Y_n

$$\mathbb{P}_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i=y}, \quad y = (x, z) \in E^2 = \mathbb{Z}^2.$$

Par conséquent, d'après (2.22), la fonction de contraste φ_n est égale

$$\varphi_n(\theta) = \mathbb{P}_n g_{\theta}(\cdot) \quad (2.26)$$

et par suite

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \mathbb{P}_n g_{\theta}(\cdot). \quad (2.27)$$

Lemme 2.4.1. *Pour tout $\theta \in \Theta$, nous avons*

1. $|f(\cdot; \theta)|^2 \in L^1(\mu_{\theta_0})$.
2. $g_{\theta}(\cdot) \in L^1(\sigma_{\theta_0})$.

Preuve. Dans la preuve suivante, nous notons par c une constante générique dont la valeur exacte peut changer pendant le développement mathématique.

Rappelons que pour tout $a \in \mathbb{R}$, nous avons $|\langle a \rangle| = \langle |a| \rangle \leq |a| + \frac{1}{2}$.

Comme Θ est compact, pour tout $x \in E = \mathbb{Z}$, nous avons

$$\begin{aligned} |f(x; \theta)| &= |\langle \alpha x + \lambda \rangle| = \langle |\alpha x + \lambda| \rangle \\ &\leq |\alpha x + \lambda| + \frac{1}{2} \leq |\alpha||x| + |\lambda| + \frac{1}{2} \\ &\leq c(1 + |x|). \end{aligned}$$

Par conséquent, $|f(x; \theta)|^2 \leq c(1 + |x|^2)$. Ainsi, comme $k \geq 2$ et $\mu_{\theta_0}(|\cdot|^k) < \infty$, la première conclusion est vérifiée.

Maintenant, pour tout $y = (x, z) \in E^2 = \mathbb{Z}^2$, nous avons

$$\begin{aligned} |z - f(x; \theta)| &= |z - \langle \alpha x + \lambda \rangle| \leq c(1 + (|x| + |z|)) \\ &= c(1 + \|y\|_1). \end{aligned}$$

Ainsi, $g_{\theta}(y) = |z - f(x; \theta)|^2 \leq c(1 + \|y\|_1^2)$. Ainsi, comme $k \geq 2$ et $\sigma_{\theta_0}(\|\cdot\|^k) < \infty$, pour toute norme sur E^2 , la deuxième conclusion est vérifiée. \square

Soit $\mathcal{G} = \{g_\theta, \theta \in \Theta\}$ la classe des fonctions de contraste, indexée par Θ compact.

D'après la preuve précédente, Pour tout $\theta \in \Theta$ et $y \in E^2$, nous avons

$$g_\theta(y) \leq A(y), \quad (2.28)$$

où

$$A(y) = c (1 + \|y\|^2) \text{ avec } c \text{ est une constante.} \quad (2.29)$$

La fonction A s'appelle la fonction enveloppe de la classe \mathcal{G} .

La proposition suivante nous donne la limite de la fonction de contraste φ_n , où la convergence se tient uniformément sur Θ .

Proposition 2.4.1. *Supposons que [H1] soit vérifiée. Donc, pour tout $\theta \in \Theta$,*

1. $\varphi_n(\theta) \xrightarrow{p.s.} K(\theta);$
2. $K(\theta) - K(\theta_0) = \mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right).$

De plus, nous avons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0.$$

Preuve. La première conclusion est simple à vérifier. Comme $\mathbb{P}_n \xrightarrow{p.s.} \sigma_{\theta_0}$ et, d'après le Lemme 2.4.1, $g_\theta \in L^1(\sigma_{\theta_0})$, nous obtenons

$$\varphi_n(\theta) = \mathbb{P}_n g_\theta \xrightarrow{p.s.} \sigma_{\theta_0} g_\theta = K(\theta).$$

Rappelons que $\sigma_{\theta_0} = \mu_{\theta_0} \otimes \pi_{\theta_0}$, où π_{θ_0} est la probabilité de transition de la chaîne (X_t) sous le vrai modèle. De plus,

$$\sum_{j \in E} j \pi_{\theta_0}(i, j) = \mathbb{E}[X_1 | X_0 = i] = f(i; \theta_0).$$

Ainsi, pour la deuxième conclusion, nous avons

$$\begin{aligned} K(\theta) - K(\theta_0) &= \sum_{(i,j) \in E^2} \sigma_{\theta_0}(i, j) (g_\theta((i, j)) - g_{\theta_0}((i, j))). \\ &= \sum_{(i,j) \in E^2} \sigma_{\theta_0}(i, j) (f(i; \theta)^2 - f(i; \theta_0)^2) - 2 \sum_{(i,j) \in E^2} \sigma_{\theta_0}(i, j) j (f(i; \theta) - f(i; \theta_0)). \\ &= \sum_{i \in E} \mu_{\theta_0}(i) (f(i; \theta)^2 - f(i; \theta_0)^2) - 2 \sum_{i \in E} \mu_{\theta_0}(i) (f(i; \theta) - f(i; \theta_0)) \sum_{j \in E} \pi_{\theta_0}(i, j) j. \\ &= \sum_{i \in E} \mu_{\theta_0}(i) (f(i; \theta)^2 - f(i; \theta_0)^2) - 2 \sum_{i \in E} \mu_{\theta_0}(i) (f(i; \theta) - f(i; \theta_0)) f(i; \theta_0). \\ &= \sum_{i \in E} \mu_{\theta_0}(i) (f(i; \theta) - f(i; \theta_0))^2. \\ &= \mu_{\theta_0} \left((f(\cdot; \theta) - f(\cdot; \theta_0))^2 \right). \end{aligned}$$

Finalement, notre but est de montrer la convergence uniforme sur Θ , c.à.d.

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0.$$

Soit $q > 0$ fixé. Rappelons que la fonction A , définie par (2.29), est la fonction enveloppe de la classe des fonctions de contraste $\mathcal{G} = \{g_\theta, \theta \in \Theta\}$.

Donc, pour tout $\theta \in \Theta$, nous trouvons

$$\begin{aligned} |\varphi_n(\theta) - K(\theta)| &= |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| \\ &= |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{\|y\| < q} + (\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{\|y\| > q}| \\ &\leq |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{\|y\| < q}| + \mathbb{P}_n(|g_\theta| \mathbb{1}_{\|y\| > q}) + \sigma_{\theta_0}(|g_\theta| \mathbb{1}_{\|y\| > q}) \\ &\leq |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{\|y\| < q}| + \mathbb{P}_n(A(y) \mathbb{1}_{\|y\| > q}) + \sigma_{\theta_0}(A(y) \mathbb{1}_{\|y\| > q}). \end{aligned}$$

Nous notons, pour tout $y \in E^2$, $p_y = \sigma_{\theta_0}(y)$ et $p_y^n = \mathbb{P}_n(y)$. De plus,

$$\begin{aligned} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{\|y\| < q}| &= \left| \sum_{\|y\| < q} g_\theta(y)(p_y^n - p_y) \right| \\ &\leq \sum_{\|y\| < q} A(y) |p_y^n - p_y|. \end{aligned}$$

Cela signifie que

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| = \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| \leq \sum_{\|y\| < q} A(y) |p_y^n - p_y| + (\mathbb{P}_n + \sigma_{\theta_0})(A(y) \mathbb{1}_{\|y\| > q}).$$

Nous avons $p_y^n \rightarrow p_y$ presque sûrement, quand $n \rightarrow \infty$. Par suite, la somme finie

$$\sum_{\|y\| < q} A(y) |p_y^n - p_y| \rightarrow 0, \text{ p.s.}$$

Ainsi,

$$\limsup_n \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| \leq 0 + 2\sigma_{\theta_0}(A(y) \mathbb{1}_{\|y\| > q}), \text{ p.s.}$$

En prenant $q \nearrow \infty$, nous obtenons presque sûrement,

$$\limsup_n \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| = 0.$$

□

Dans les paragraphes suivants nous traitons le problème d'identifiabilité du modèle du à l'opérateur d'arrondi. Nous constatons deux cas dépendants du coefficient de régression α_0 .

Notamment, si α_0 est rationnel, il y a une manque d'identifiabilité autour de λ_0 .

Par la suite, nous définissons sur l'espace des paramètres Θ la distance d par

$$d(\theta, \theta') = \max \left\{ |\alpha - \alpha'|, |\lambda - \lambda'| \right\}, \forall \theta, \theta' \in \Theta. \quad (2.30)$$

2.4.1 La consistance forte de l'estimateur des moindres carrés quand le coefficient de régression α_0 est irrationnel

Rappelons que $\theta_0 = (\alpha_0, \lambda_0)$ est la vraie valeur du paramètre θ du modèle RINAR(1). La proposition suivante aborde la question d'identifiabilité des paramètres de ce dernier pour le cas où α_0 est irrationnel.

Proposition 2.4.2. *Supposons que [H1] soit vérifiée. Si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors pour tout $\theta \in \Theta$*

$$f(x; \theta) = f(x; \theta_0), \forall x \in E \iff \theta = \theta_0.$$

Preuve. Notre but est de montrer que si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors

$$\langle \alpha x + \lambda \rangle = \langle \alpha_0 x + \lambda_0 \rangle, \forall x \in E \Rightarrow \alpha = \alpha_0 \text{ et } \lambda = \lambda_0.$$

L'idée de la preuve est de vérifier que si $\theta \neq \theta_0$ alors il existe $x_0 \in E$ tel que

$$f(x_0; \theta) \neq f(x_0; \theta_0).$$

Notons que pour tout $a, b \in \mathbb{R}$, nous avons

$$||\langle a \rangle - \langle b \rangle| - |a - b|| \leq |(\langle a \rangle - \langle b \rangle) - (a - b)| \leq 1,$$

et par suite

$$|\langle a \rangle - \langle b \rangle| \geq |a - b| - 1.$$

Comme Θ est compact, il existe $B > 0$ tel que $|\lambda| \leq B$. Par conséquent, pour tout $x \in E = \mathbb{Z}$, nous obtenons

$$\begin{aligned} |f(x; \theta) - f(x; \theta_0)| &= |\langle \alpha x + \lambda \rangle - \langle \alpha_0 x + \lambda_0 \rangle| \\ &\geq |(\alpha - \alpha_0) x + (\lambda - \lambda_0)| - 1 \\ &\geq |\alpha - \alpha_0| |x| - |\lambda - \lambda_0| - 1 \\ &\geq |\alpha - \alpha_0| |x| - 2B - 1. \end{aligned}$$

Si $\alpha \neq \alpha_0$, alors il existe $x_0 \in E = \mathbb{Z}$ tel que $|f(x_0; \theta) - f(x_0; \theta_0)| > 0$,

c'est une contradiction. Ensuite, pour $x_0 = 0$, $f(0; \theta) = f(0; \theta_0)$ implique que $\langle \lambda \rangle = \langle \lambda_0 \rangle$.

Ainsi, sans utiliser le fait que $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, nous avons

$$\alpha = \alpha_0 \text{ et } \langle \lambda \rangle = \langle \lambda_0 \rangle.$$

Rappelons que $\{a\}$ représente la partie fractionnaire de $a \in \mathbb{R}$. Maintenant, nous supposons sans perte de généralité, $\alpha_0 > 0$, $\lambda \geq \lambda_0 \geq 0$ et $\{\lambda\}, \{\lambda_0\} \in [0, \frac{1}{2}[$.

Il se découle,

$$|\lambda - \lambda_0| = \lambda - \lambda_0 = \{\lambda\} - \{\lambda_0\}$$

et

$$f(x; \theta) = f(x; \theta_0), \forall x \in E \Rightarrow \langle \{\alpha_0 x\} + \{\lambda\} \rangle = \langle \{\alpha_0 x\} + \{\lambda_0\} \rangle, \forall x \in E. \quad (2.31)$$

Comme $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, $(\{\alpha_0 x\})_{x \in \mathbb{Z}}$ est dense dans $[0, 1[$. Si $\{\lambda\} \neq \{\lambda_0\}$, alors il existe $x_0 \in \mathbb{Z}$ tel que

$$\{\alpha_0 x_0\} + \{\lambda_0\} < \frac{1}{2} \text{ et } \{\alpha_0 x_0\} + \{\lambda\} \geq \frac{1}{2}.$$

C'est une contradiction avec (2.31). \square

Lemme 2.4.2. *Soit $x_0 \in \mathbb{N}^*$ fixé, la fonction $\alpha \longrightarrow \{\alpha x_0\}$, définie sur \mathbb{R} et à valeurs dans $[0, 1[$, est continue en α_0 si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$.*

Preuve. $x_0 \in \mathbb{N}^*$ fixé. Nous avons, $\{\alpha x_0\} = |\alpha x_0| - [\alpha x_0]$, où $[\cdot]$ est la fonction partie entière. Alors la fonction $\{\cdot\} : \mathbb{R} \rightarrow [0, 1[$, définie par $\alpha \rightarrow \{\alpha x_0\}$, est discontinue en α si $\alpha x_0 \in \mathbb{Z}$. Par conséquent, si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors $\alpha \longrightarrow \{\alpha x_0\}$ est continue en α_0 . \square

Quelques rappels sont nécessaires. Supposons que [H1] soit vérifiée, alors d'après la Proposition 2.4.1 nous avons,

$$K(\theta) - K(\theta_0) = \sigma_{\theta_0} [g_{\theta}(\cdot) - g_{\theta_0}(\cdot)] = \mu_{\theta_0} [f(\cdot; \theta) - f(\cdot; \theta_0)]^2, \quad (2.32)$$

où

$$g_{\theta}(y) = [z - f(x; \theta)]^2, \forall y = (x, z) \in E^2,$$

et

$$f(x; \theta) = \langle \alpha x + \lambda \rangle, \forall x \in E \text{ et } \theta = (\alpha, \lambda) \in \Theta.$$

Proposition 2.4.3. *Supposons que [H1] soit vérifiée. Si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$, alors pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_{\varepsilon}} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_{\varepsilon} = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}$.

Preuve. Rappelons que le support de la mesure stationnaire du processus (X_t) est $E = \mathbb{Z}$. D'autre part, notons que les deux composantes du paramètre θ ne possèdent pas les mêmes rôles. Pour cela, nous distinguons trois situations pour l'événement Θ_ε .

Ainsi, notons que $\Theta_\varepsilon = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, où

$$\Gamma_1 = \{\theta \in \Theta, |\alpha - \alpha_0| \geq \varepsilon\}, \quad (2.33)$$

$$\Gamma_2 = \{\theta \in \Theta, |\alpha - \alpha_0| < \varepsilon, |\lambda - \lambda_0| \geq \varepsilon, \langle \lambda \rangle \neq \langle \lambda_0 \rangle\}, \quad (2.34)$$

et

$$\Gamma_3 = \{\theta \in \Theta, |\alpha - \alpha_0| < \varepsilon, |\lambda - \lambda_0| \geq \varepsilon, \langle \lambda \rangle = \langle \lambda_0 \rangle\}. \quad (2.35)$$

Nous allons montrer

$$\inf_{\theta \in \Gamma_i} |K(\theta) - K(\theta_0)| > 0, \quad i = 1, 2, 3.$$

L'idée de la preuve est basée sur l'équation (2.32). Alors, le but est de trouver $x_0 \in E$ tel que

$$|f(x_0; \theta) - f(x_0; \theta_0)| > 0, \text{ uniformément sur } \Gamma_i, \quad i = 1, 2, 3.$$

1. Nous considérons le premier cas, $\theta = (\alpha, \lambda) \in \Gamma_1$. Comme Θ est compact, il existe $A > 0$ et $B > 0$ tels que $|\alpha| \leq A < 1$ et $|\lambda| \leq B$. Par suite,

$$|f(x; \theta) - f(x; \theta_0)| \geq |\alpha - \alpha_0||x| - 2B - 1 \geq \varepsilon|x| - 2B - 1.$$

Par conséquent, il existe $x_0 > 0$ tel que $\forall |x| \geq x_0$, nous avons

$$|f(x; \theta) - f(x; \theta_0)| \geq \frac{\varepsilon}{2}|x|, \text{ uniformément sur } \Gamma_1.$$

Il en découle, puisque le support de μ_{θ_0} n'est pas borné,

$$\inf_{\theta \in \Gamma_1} |K(\theta) - K(\theta_0)| \geq \int_{|x| \geq x_0} \left(\frac{\varepsilon}{2}x\right)^2 \mu_{\theta_0}(dx) > 0.$$

2. Pour $\theta \in \Gamma_2$, $\langle \lambda \rangle \neq \langle \lambda_0 \rangle$ donc $|\langle \lambda \rangle - \langle \lambda_0 \rangle| \geq 1$. Comme 0 appartient au support de μ_{θ_0} , nous prenons $x_0 = 0$ et par suite

$$|f(0; \theta) - f(0; \theta_0)| = |\langle \lambda \rangle - \langle \lambda_0 \rangle| \geq 1.$$

Par conséquent,

$$\inf_{\theta \in \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

3. Pour $\theta \in \Gamma_3$, nous supposons sans perte de généralité, $\alpha_0 > 0$ et $\lambda_0 \geq 0$.

Nous distinguons quatres cas dépendants de la position de la partie fractionnaire de λ_0 et λ :

(a) Cas A : $\{\lambda_0\} \in [0, \frac{1}{2}[$ et $\lambda \geq \lambda_0$.

(b) Cas B : $\{\lambda_0\} \in [\frac{1}{2}, 1[$ et $\lambda \geq \lambda_0$.

(c) Cas C : $\{\lambda_0\} \in [0, \frac{1}{2}[$ et $\lambda \leq \lambda_0$.

(d) Cas D : $\{\lambda_0\} \in [\frac{1}{2}, 1[$ et $\lambda \leq \lambda_0$.

– Cas A : $\{\lambda_0\} \in [0, \frac{1}{2}[$ et $\lambda \geq \lambda_0$. Ici, il est simple de vérifier que

$$\{\lambda\} \in \left[0, \frac{1}{2}\right] \text{ et } \lambda - \lambda_0 = \{\lambda\} - \{\lambda_0\} \geq \varepsilon. \quad (2.36)$$

Donc, le but est de trouver $x_0 \in \mathbb{N}$ tel que

$$\{\alpha_0 x_0\} + \{\lambda_0\} < \frac{1}{2}, \quad (2.37)$$

et pour α suffisamment proche de α_0 et tout λ tel que $|\lambda - \lambda_0| \geq \varepsilon$, nous trouvons

$$\{\alpha x_0\} + \{\lambda\} \geq \frac{1}{2}. \quad (2.38)$$

Soit ε suffisamment petit, tel que $\{\lambda_0\} + \varepsilon < \frac{1}{2}$. Comme $(\{\alpha_0 x\})_{x \in \mathbb{N}}$ est dense dans $[0, 1[$, il existe $x_0 \in \mathbb{N}$ tel que

$$|\{\alpha_0 x_0\} - (\frac{1}{2} - \{\lambda_0\} - \frac{\varepsilon}{2})| < \frac{\varepsilon}{4}. \quad (2.39)$$

Par suite l'inégalité (2.37) est immédiatement vérifiée. De plus, pour ce même $x_0 \in \mathbb{N}$ fixé, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue en α_0 (voir Lemme 2.4.2).

Ainsi, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$, tel que $\forall |\alpha - \alpha_0| \leq \eta$ nous trouvons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0], \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.40)$$

Ainsi, d'après (2.36) (2.39) et (2.40), nous obtenons

$$\begin{aligned} \{\alpha x_0\} + \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} + \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\} \\ &\geq -\frac{\varepsilon}{4} + \varepsilon + (\frac{1}{2} - \frac{\varepsilon}{2} - \frac{\varepsilon}{4}) = \frac{1}{2}. \end{aligned}$$

Donc, l'inégalité (2.38) est vérifiée.

Finalement, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous trouvons

$$\begin{aligned}
 f(x_0; \theta) - f(x_0; \theta_0) &= \langle \alpha x_0 + \lambda \rangle - \langle \alpha_0 x_0 + \lambda_0 \rangle, \\
 &= \langle [\alpha x_0] + \langle \lambda \rangle + \{\alpha x_0\} + \{\lambda\} \rangle - \langle [\alpha_0 x_0] + \langle \lambda_0 \rangle + \{\alpha_0 x_0\} + \{\lambda_0\} \rangle, \\
 &= [\alpha x_0] + \langle \lambda \rangle + \langle \{\alpha x_0\} + \{\lambda\} \rangle - [\alpha_0 x_0] - \langle \lambda_0 \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle, \\
 &= \langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle, \\
 &= 1.
 \end{aligned}$$

– Cas B : $\{\lambda_0\} \in [\frac{1}{2}, 1[$ et $\lambda \geq \lambda_0$. Nous distinguons deux cas :

– Si $\{\lambda\} \in [\frac{1}{2}, 1[$. Il est simple de vérifier que

$$\lambda - \lambda_0 = \{\lambda\} - \{\lambda_0\} \geq \varepsilon. \quad (2.41)$$

Ici, le but est de trouver $x_0 \in \mathbb{N}$ tel que

$$\{\alpha_0 x_0\} + \{\lambda_0\} < \frac{3}{2}, \quad (2.42)$$

et pour α suffisamment proche de α_0 et tout λ tel que $|\lambda - \lambda_0| \geq \varepsilon$

$$\{\alpha x_0\} + \{\lambda\} \geq \frac{3}{2}. \quad (2.43)$$

Soit ε suffisamment petit, tel que $\{\lambda_0\} + \varepsilon < 1$. Comme $(\{\alpha_0 x\})_{x \in \mathbb{N}}$ est dense dans $[0, 1[$, il existe $x_0 \in \mathbb{N}$ tel que

$$|\{\alpha_0 x_0\} - (\frac{3}{2} - \{\lambda_0\} - \frac{\varepsilon}{2})| < \frac{\varepsilon}{4}. \quad (2.44)$$

Par suite, l'inégalité (2.42) est immédiatement vérifiée. D'autre part, pour ce même $x_0 \in \mathbb{N}$ fixé, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue en α_0 (voir Lemme 2.4.2). Ainsi, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$, tel que $\forall |\alpha - \alpha_0| \leq \eta$ nous trouvons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0], \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.45)$$

Par conséquent, d'après (2.41) (2.44) et (2.45), nous obtenons

$$\begin{aligned}
 \{\alpha x_0\} + \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} + \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}. \\
 &\geq -\frac{\varepsilon}{4} + \varepsilon + (\frac{3}{2} - \frac{\varepsilon}{2} - \frac{\varepsilon}{4}) = \frac{3}{2}.
 \end{aligned}$$

Donc, l'inégalité (2.43) est vérifiée.

Ainsi, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous trouvons

$$|f(x_0; \theta) - f(x_0; \theta_0)| = |\langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle| = 1.$$

– Si $\{\lambda\} \in [0, \frac{1}{2}[$, il en découle

$$\lambda - \lambda_0 = 1 + \{\lambda\} - \{\lambda_0\} \geq \varepsilon \text{ et } [\lambda] = [\lambda_0] + 1, \quad (2.46)$$

où $[\cdot]$ est la partie entière. Pour le même x_0 du cas précédent, nous retrouvons (2.44) et (2.45). Finalement, d'après (2.46), nous obtenons

$$\begin{aligned} \{\alpha x_0\} + \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} + \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}, \\ &\geq -\frac{\varepsilon}{4} + \varepsilon - 1 + \left(\frac{3}{2} - \frac{\varepsilon}{2} - \frac{\varepsilon}{4}\right) = \frac{1}{2}. \end{aligned}$$

Par conséquent,

$$\begin{aligned} f(x_0; \theta) - f(x_0; \theta_0) &= \langle \alpha x_0 + \lambda \rangle - \langle \alpha_0 x_0 + \lambda_0 \rangle, \\ &= \langle [\alpha x_0] + [\lambda] + \{\alpha x_0\} + \{\lambda\} \rangle \\ &\quad - \langle [\alpha_0 x_0] + [\lambda_0] + \{\alpha_0 x_0\} + \{\lambda_0\} \rangle, \\ &= ([\alpha x_0] - [\alpha_0 x_0]) + ([\lambda] - [\lambda_0]) \\ &\quad + (\langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle), \\ &= 1. \end{aligned}$$

– Cas C : Nous supposons que $\{\lambda_0\} \in [0, \frac{1}{2}[$ et $\lambda \leq \lambda_0$. Nous distinguons deux cas.

– Le cadre le plus intéressant à examiner est quand $\langle \lambda_0 \rangle = \langle \lambda \rangle = 0$.

Par conséquent, nous trouvons,

$$\lambda_0 = \{\lambda_0\} \in \left[0, \frac{1}{2}\right[\text{ et } \lambda = \begin{cases} \{\lambda\} & \text{si } 0 \leq \lambda \leq \lambda_0 < \frac{1}{2}, \\ -\{\lambda\} & \text{si } -\frac{1}{2} < \lambda < 0 \leq \lambda_0 < \frac{1}{2}. \end{cases}$$

Par conséquent, $\{\lambda\} \in [0, \frac{1}{2}[$ et

$$|\lambda - \lambda_0| = \begin{cases} \{\lambda_0\} - \{\lambda\} & \text{si } \lambda \geq 0, \\ \{\lambda_0\} + \{\lambda\} & \text{si } \lambda < 0. \end{cases}$$

Ainsi, nous distinguons deux cas :

– Supposons que $0 \leq \lambda \leq \lambda_0 < \frac{1}{2}$. Ici,

$$\{\lambda_0\} \text{ et } \{\lambda\} \in \left[0, \frac{1}{2}\right[\quad (2.47)$$

et

$$|\lambda - \lambda_0| = \lambda_0 - \lambda = \{\lambda_0\} - \{\lambda\} \geq \varepsilon. \quad (2.48)$$

Notre but est de trouver $x_0 \in \mathbb{N}$ tel que

$$\{\alpha_0 x_0\} + \{\lambda_0\} \geq \frac{1}{2} \quad (2.49)$$

et pour α suffisamment proche de α_0 et tout λ tel que $|\lambda - \lambda_0| \geq \varepsilon$

$$\{\alpha x_0\} + \{\lambda\} < \frac{1}{2}. \quad (2.50)$$

Comme $(\{\alpha_0 x\})_{x \in \mathbb{N}}$ est dense dans $[0, 1[$, il existe $x_0 \in \mathbb{N}$ tel que

$$|\{\alpha_0 x_0\} - (\frac{1}{2} - \{\lambda_0\} + \frac{\varepsilon}{2})| < \frac{\varepsilon}{4}. \quad (2.51)$$

Par conséquent,

$$\frac{1}{2} + \frac{\varepsilon}{4} < \{\alpha_0 x_0\} + \{\lambda_0\} < \frac{1}{2} + \frac{3\varepsilon}{4}$$

Notons que, pour ce même $x_0 \in \mathbb{N}$ fixé, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue en α_0 (voir Lemme 2.4.2). Par conséquent, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$, nous trouvons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0], \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.52)$$

D'après (2.47) (2.48) (2.51) et (2.52), nous obtenons

$$\begin{aligned} \{\alpha x_0\} + \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} + \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}. \\ &< \frac{\varepsilon}{4} - \varepsilon + \frac{1}{2} + \frac{3\varepsilon}{4} = \frac{1}{2}. \end{aligned}$$

Finalement, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous obtenons

$$\begin{aligned} |f(x_0; \theta) - f(x_0; \theta_0)| &= |\langle \alpha x_0 + \lambda \rangle - \langle \alpha_0 x_0 + \lambda_0 \rangle| \\ &= |\langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle| = 1. \end{aligned}$$

– Maintenant, nous supposons que $-\frac{1}{2} < \lambda < 0 \leq \lambda_0 < \frac{1}{2}$. Nous avons,

$$\{\lambda_0\}, \{\lambda\} \in \left[0, \frac{1}{2}\right] \text{ et } \lambda = -\{\lambda\}. \quad (2.53)$$

Par conséquent, nous trouvons

$$|\lambda - \lambda_0| = \lambda_0 - \lambda = \{\lambda_0\} + \{\lambda\} \geq \varepsilon. \quad (2.54)$$

Ensuite, nous prenons le même x_0 du cas précédent. Donc,

$$\frac{1}{2} + \frac{\varepsilon}{4} < \{\alpha_0 x_0\} + \{\lambda_0\} < \frac{1}{2} + \frac{3\varepsilon}{4} \quad (2.55)$$

De plus, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$, nous avons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0], \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.56)$$

D'après (2.53) (2.54) (2.55) et (2.56), il en découle

$$\begin{aligned} \{\alpha x_0\} - \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} - \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}. \\ &< \frac{\varepsilon}{4} - \varepsilon + \frac{1}{2} + \frac{3\varepsilon}{4} = \frac{1}{2}. \end{aligned}$$

Finalement, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous obtenons

$$\begin{aligned} |f(x_0; \theta) - f(x_0; \theta_0)| &= |\langle \alpha x_0 + \lambda \rangle - \langle \alpha_0 x_0 + \lambda_0 \rangle| \\ &= |\langle \{\alpha x_0\} - \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle| = 1. \end{aligned}$$

- Nous supposons que $\{\lambda_0\} \in [0, \frac{1}{2}[$, $\lambda \leq \lambda_0$ et $\langle \lambda_0 \rangle = \langle \lambda \rangle \geq 1$.

Nous distinguons deux cas :

- Si $\{\lambda\} \in [0, \frac{1}{2}[$. Alors, $0 \leq \{\lambda\} \leq \{\lambda_0\} < \frac{1}{2}$ et $|\lambda - \lambda_0| = \{\lambda_0\} - \{\lambda\} \geq \varepsilon$.

L'étude de ce cas est similaire à celle du cas où $0 \leq \lambda \leq \lambda_0 < \frac{1}{2}$ traitée précédemment.

- Si $\{\lambda\} \in [\frac{1}{2}, 1[$. Alors, nous avons

$$[\lambda_0] - [\lambda] = 1 \text{ et par conséquent } |\lambda - \lambda_0| = 1 + \{\lambda_0\} - \{\lambda\} \geq \varepsilon. \quad (2.57)$$

Ici, par les mêmes arguments utilisés précédemment, il existe $x_0 \in \mathbb{N}$ tel que

$$|\{\alpha_0 x_0\} - (\frac{1}{2} - \{\lambda_0\} + \frac{\varepsilon}{2})| < \frac{\varepsilon}{4}. \quad (2.58)$$

De plus, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$, nous avons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0]. \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.59)$$

En utilisant (2.57) (2.58) et (2.59), nous obtenons

$$\begin{aligned} \{\alpha x_0\} - \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} - \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}. \\ &< \frac{\varepsilon}{4} + 1 - \varepsilon + \frac{1}{2} + \frac{3\varepsilon}{4} = \frac{3}{2}. \end{aligned}$$

Donc, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous obtenons

$$|f(x_0; \theta) - f(x_0; \theta_0)| = |[\lambda] - [\lambda_0]| = 1.$$

– Cas D : $\{\lambda_0\} \in [\frac{1}{2}, 1[$ et $\lambda \leq \lambda_0$. Il est simple de vérifier que

$$\{\lambda\} \in \left[\frac{1}{2}, 1\right[\text{ et } |\lambda - \lambda_0| = \lambda_0 - \lambda = \{\lambda_0\} - \{\lambda\} \geq \varepsilon \quad (2.60)$$

Ainsi, en utilisant les mêmes arguments précédents, il existe $x_0 \in \mathbb{N}$ tel que

$$|\{\alpha_0 x_0\} - (\frac{3}{2} - \{\lambda_0\} + \frac{\varepsilon}{2})| < \frac{\varepsilon}{4}. \quad (2.61)$$

Il en découle,

$$\{\alpha_0 x_0\} + \{\lambda_0\} \geq \frac{3}{2}$$

De plus, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$, nous avons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0]. \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \frac{\varepsilon}{4}. \end{cases} \quad (2.62)$$

Par conséquent, d'après (2.60) (2.61) et (2.62), nous trouvons

$$\begin{aligned} \{\alpha x_0\} + \{\lambda\} &= \{\alpha x_0\} - \{\alpha_0 x_0\} + \{\lambda\} - \{\lambda_0\} + \{\alpha_0 x_0\} + \{\lambda_0\}. \\ &< \frac{\varepsilon}{4} - \varepsilon + \frac{3}{2} + \frac{3\varepsilon}{4} = \frac{3}{2}. \end{aligned}$$

Finalement, $\forall \theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous obtenons

$$|f(x_0; \theta) - f(x_0; \theta_0)| = |\langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle| = 1.$$

Alors, nous déduisons qu'il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ et une constante $e_0 > 0$, tel que uniformément sur Γ'_3 nous avons $|f(x_0; \theta) - f(x_0; \theta_0)| \geq e_0 > 0$.

Pour $\theta \in \Gamma''_3 = \Gamma_3 \setminus \Gamma'_3 = \{\theta \in \Gamma_3 : |\alpha - \alpha_0| > \eta\}$ les mêmes argument pour Γ_1 peuvent être appliqués ici et par suite il existe $d_0 > 0$ tel que

$$\inf_{\theta \in \Gamma''_3} |K(\theta) - K(\theta_0)| \geq d_0 > 0.$$

□

Théorème 2.4.1. *Supposons que*

1. α_0 , la vraie valeur de α , est un nombre irrationnel appartenant à l'intervalle $] -1, 1[$;
2. Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible ;
3. Pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;
4. L'espace des paramètres Θ , est un compact sous-ensemble $] -1, 1[\times \mathbb{R}$.

Alors, l'estimateur des moindres carrés $\hat{\theta}_n$ de θ_0 est fortement consistant, c.à.d.

$$\hat{\theta}_n \xrightarrow{p.s.} \theta_0.$$

Preuve. Rappelons que, D'après la Proposition 2.4.1 et 2.4.3, nous avons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0$$

et

$$\inf_{\theta \in \Theta_\varepsilon} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}$ et d est la distance sur l'espace des paramètres Θ définie par

$$d(\alpha, \alpha') = |\alpha - \alpha'|, \forall \alpha, \alpha' \in \Theta.$$

Ainsi, en utilisant le Théorème 5.7 de Van Der Vaart (1998, [61]), nous obtenons que $\hat{\theta}_n \rightarrow \theta_0$ presque sûrement. \square

2.4.2 La consistance forte de l'estimateur des moindres carrés quand le coefficient de régression α_0 est rationnel

Tout d'abord, nous donnons quelques explications nécessaires pour comprendre le problème d'identifiabilité du modèle RINAR(1), quand α_0 est rationnel.

Soit $\alpha_0 = \frac{p}{q} < 1$, où $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$, p et q sont premiers entre eux ($p \wedge q = 1$).

Rappelons que $\{a\}$ représente la partie fractionnaire de $a \in \mathbb{R}$.

Pour tout p tel que $|p| < |q|$ et pour toute valeur de $x \in \mathbb{Z}$, la suite $(\{\alpha_0 x\})_{x \in \mathbb{Z}}$ ne prend qu'un nombre fini de valeurs : $\left\{0, \frac{1}{q}, \dots, \frac{q-1}{q}\right\}$, c.à.d.

$$\text{Im}[(\{\alpha_0 x\})_{x \in \mathbb{Z}}] = \left\{0, \frac{1}{q}, \dots, \frac{q-1}{q}\right\}.$$

Nous distinguons deux cas :

- Si q est pair, alors $\frac{1}{2} \in \text{Im}[(\{\alpha_0 x\})_{x \in \mathbb{Z}}]$. Par conséquent, la subdivision de l'intervalle $[0, 1[$ en q intervalles de taille $\frac{1}{q}$, montre que $\frac{1}{2}$ est un point frontière de deux intervalles $\left[\frac{\frac{q}{2}-1}{q}, \frac{1}{2}\right[$ et $\left[\frac{1}{2}, \frac{\frac{q}{2}+1}{q}\right[$. En d'autres termes, il y a $\frac{q}{2}$ intervalles avant et après $\frac{1}{2}$.

Maintenant, supposons que a et b appartiennent à l'un de q intervalles de la subdivision, tels que $a \neq b$ (voir Figure 2.4), nous obtenons

$$\forall x \in \mathbb{Z}, \langle \{\alpha_0 x\} + a \rangle = \langle \{\alpha_0 x\} + b \rangle \text{ et } \langle \{\alpha_0 x\} - a \rangle = \langle \{\alpha_0 x\} - b \rangle. \quad (2.63)$$

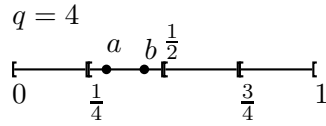


FIG. 2.4 – La subdivision de l'intervalle $[0, 1[$ en 4 intervalles de taille $\frac{1}{4}$

- Si q est impair, alors $\frac{1}{2} \notin \left\{0, \frac{1}{q}, \dots, \frac{q-1}{q}\right\}$. Donc, la subdivision de l'intervalle $[0, 1[$ en q intervalles de taille $\frac{1}{q}$, montre qu'ici $\frac{1}{2}$ est le centre de l'intervalle $\left[\frac{\langle \frac{q}{2} \rangle - 1}{q}, \frac{\langle \frac{q}{2} \rangle}{q}\right]$. Ainsi, l'équation (2.63) n'est plus vérifiée. Par exemple, pour $q = 5$ (voir Figure 2.5), soient $a, b \in \left[\frac{1}{5}, \frac{2}{5}\right[$ tels que $a = 0.23$ et $b = 0.36$. Alors, pour $x_0 = 1$ et $p = 1$, nous avons

$$\langle \{\alpha_0 x_0\} + a \rangle = \langle \frac{1}{5} + 0.23 \rangle = 0 \neq \langle \{\alpha_0 x_0\} + b \rangle = \langle \frac{1}{5} + 0.36 \rangle = 1.$$

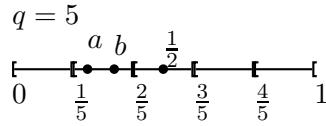
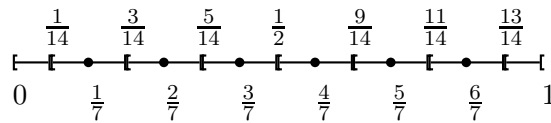


FIG. 2.5 – La subdivision de l'intervalle $[0, 1[$ en 5 intervalles de taille $\frac{1}{5}$

Maintenant, nous introduisons une nouvelle subdivision de l'intervalle $[0, 1[$ en $q + 1$ intervalles. Le premier est $\left[0, \frac{1}{2q}\right]$, les $q - 1$ prochaines intervalles sont $\left[\frac{i}{2q}, \frac{i+2}{2q}\right]$ pour $i = 1, 3, \dots, q, \dots, 2q - 3$ et le dernier est $\left[\frac{2q-1}{2q}, 1\right]$. Pour cette subdivision, $\frac{1}{2}$ est un point frontière de deux intervalles $\left[\frac{q-2}{2q}, \frac{1}{2}\right]$ et $\left[\frac{1}{2}, \frac{q+2}{2q}\right]$. En d'autres termes, il y a $\langle \frac{q}{2} \rangle$ intervalles avant et après $\frac{1}{2}$ dont un de taille $\frac{1}{2q}$ et les autres sont de taille $\frac{1}{q}$ (voir Figure 2.6). De plus, supposons que a et b appartiennent à l'un des $q + 1$ intervalles de cette nouvelle subdivision, tels que $a \neq b$, nous obtenons

$$\forall x \in \mathbb{Z}, \langle \{\alpha_0 x\} + a \rangle = \langle \{\alpha_0 x\} + b \rangle \text{ et } \langle \{\alpha_0 x\} - a \rangle = \langle \{\alpha_0 x\} - b \rangle. \quad (2.64)$$

FIG. 2.6 – La nouvelle subdivision de l'intervalle $[0, 1[$, avec $q = 7$

Notons que, l'équation (2.63) (resp. (2.64)) doit être vérifiée dans le cas où q est pair (resp. impair), pour prouver la proposition suivante.

Proposition 2.4.4. *Supposons que [H1] soit vérifiée. Si $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$, p et q sont premiers entre eux ($p \wedge q = 1$), alors*

$$f(x; \theta) = f(x; \theta_0), \quad \forall x \in E \iff \alpha = \alpha_0 \text{ et } \lambda \in I_0,$$

où I_0 est un intervalle de taille $\frac{1}{q}$ autour de λ_0 tel que

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle\}.$$

La localisation de I_0 dépend de la parité de q et de la position de $\{\lambda_0\}$, la partie fractionnaire de λ_0 , sur l'intervalle $[0, 1[$. Rappelons que d'après la discussion précédente, la parité ou l'impairité de q impose deux différentes subdivision de $[0, 1[$.

Ainsi, nous distinguons deux cas :

- Supposons que q est pair. Il existe $k_0 \in \{0, 1, \dots, q-1\}$ tel que

$$\{\lambda_0\} \in i_0 = \left[\frac{k_0}{q}, \frac{k_0+1}{q} \right[\quad (2.65)$$

Ainsi, nous définissons

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle \text{ et } \{\lambda\} \in i_0\}. \quad (2.66)$$

Par exemple, soit $q = 4$. Supposons $\lambda_0 = 2.52$, alors $\langle \lambda_0 \rangle = 3$ et $\{\lambda_0\} = 0.52$.

Par suite, nous trouvons

$$i_0 = \left[\frac{1}{2}, \frac{3}{4} \right[\text{ et } I_0 = \{\lambda, \langle \lambda \rangle = 3 \text{ et } \{\lambda\} \in i_0\} = [2.5, 2.75[.$$

Notons que si $\lambda_0 = -2.52$, alors

$$I_0 = \{\lambda, \langle \lambda \rangle = -3 \text{ et } \{\lambda\} \in i_0\} =]-2.75, -2.5].$$

- Supposons que q est impair. Nous remarquons trois sous-cas :

– Si $\{\lambda_0\} \in \left[0, \frac{1}{2q}\right]$, alors I_0 est défini par

$$I_0 = \begin{cases} \left[[\lambda_0] - 1 + \frac{2q-1}{2q}, [\lambda_0] + \frac{1}{2q} \right], & \text{si } [\lambda_0] \geq 1, \\ \left[-\frac{1}{2q}, \frac{1}{2q} \right], & \text{si } [\lambda_0] = 0, \\ \left[[\lambda_0] - \frac{1}{2q}, [\lambda_0] + 1 - \frac{2q-1}{2q} \right], & \text{si } [\lambda_0] \leq -1. \end{cases} \quad (2.67)$$

Par exemple, soit $q = 5$. Supposons $\lambda_0 = 1.08$, alors $\{\lambda_0\} = 0.08$ et $[\lambda_0] = 1$.

Par suite, nous trouvons

$$\{\lambda_0\} \in \left[0, \frac{1}{10}\right] \text{ et } I_0 = [0.9, 1.1[.$$

Notons que si $\lambda_0 = -1.08$, alors

$$I_0 =]-1.1, -0.9].$$

– Si $\{\lambda_0\} \in \left[\frac{1}{2q}, \frac{2q-1}{2q}\right]$, alors il existe $k_0 \in \{1, 3, \dots, q, \dots, 2q-3\}$ tel que

$$\{\lambda_0\} \in i_0 = \left[\frac{k_0}{2q}, \frac{k_0+2}{2q}\right]. \quad (2.68)$$

Ainsi, nous définissons

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle \text{ et } \{\lambda\} \in i_0\}. \quad (2.69)$$

Par exemple, soit $q = 5$. Supposons $\lambda_0 = 4.49$, alors $\langle \lambda_0 \rangle = 4$ et $\{\lambda_0\} = 0.49$.

Par suite, nous trouvons

$$i_0 = \left[\frac{3}{10}, \frac{1}{2}\right] \text{ et } I_0 = \{\lambda, \langle \lambda \rangle = 4 \text{ et } \{\lambda\} \in i_0\} = [4.3, 4.5[.$$

Notons que si $\lambda_0 = -4.49$, alors

$$I_0 = \{\lambda, \langle \lambda \rangle = -4 \text{ et } \{\lambda\} \in i_0\} =]-4.5, -4.3].$$

– Si $\{\lambda_0\} \in \left[\frac{2q-1}{2q}, 1\right]$, alors I_0 est défini par

$$I_0 = \begin{cases} \left[[\lambda_0] + \frac{2q-1}{2q}, [\lambda_0] + 1 + \frac{1}{2q} \right], & \text{si } \lambda_0 > 0, \\ \left[[\lambda_0] - 1 - \frac{1}{2q}, [\lambda_0] - \frac{2q-1}{2q} \right], & \text{si } \lambda_0 < 0. \end{cases} \quad (2.70)$$

Par exemple, soit $q = 5$. Supposons $\lambda_0 = 8.98$, alors $\{\lambda_0\} = 0.98$, $\langle \lambda_0 \rangle = 9$ et $[\lambda_0] = 8$. Par suite, nous trouvons

$$\{\lambda_0\} \in \left[\frac{9}{10}, 1 \right[\text{ et } I_0 = [8.9, 9.1[.$$

Notons que si $\lambda_0 = -8.98$, alors

$$I_0 =]-9.1, -8.9].$$

Preuve. Soit

$$A_0 = \left\{ \theta = (\alpha, \lambda) \in \Theta, \alpha = \alpha_0 = \frac{p}{q} \text{ et } \lambda \in I_0 \right\},$$

où $I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle \text{ et } \{\lambda\} \in i_0\}$. Rappelons que λ_0 la vraie valeur du paramètre λ appartient à I_0 et i_0 est défini en fonction de la parité de q et la position de la partie fractionnaire de λ_0 sur l'intervalle $[0, 1[$.

Notons que pour tout $a \in \mathbb{R}$, nous avons $a = [a] + s(a)\{a\}$, où $[\cdot]$ représente la fonction partie entière et $s(\cdot)$ est la fonction signe.

Soit $\theta \in A_0$. Il en découle, $\{\lambda\}$ et $\{\lambda_0\} \in i_0$, $s(\lambda) = s(\lambda_0)$ et $[\lambda] = [\lambda_0]$.

Par conséquent, pour tout $x \in \mathbb{Z}$, nous avons

$$\begin{aligned} f(x; \theta) - f(x; \theta_0) &= \langle \alpha_0 x + \lambda \rangle - \langle \alpha_0 x + \lambda_0 \rangle \\ &= \langle s(\alpha_0 x) \{ \alpha_0 x \} + s(\lambda_0) \{ \lambda \} \rangle - \langle s(\alpha_0 x) \{ \alpha_0 x \} + s(\lambda_0) \{ \lambda_0 \} \rangle. \end{aligned}$$

Nous distinguons quatre cas :

- Supposons que $s(\alpha_0 x) = 1$ et $s(\lambda_0) = 1$. Ainsi, pour tout $x \in \mathbb{Z}$, nous obtenons

$$f(x; \theta) - f(x; \theta_0) = \langle \{ \alpha_0 x \} + \{ \lambda \} \rangle - \langle \{ \alpha_0 x \} + \{ \lambda_0 \} \rangle.$$

Par suite, d'après (2.63) (resp. (2.64)) dans le cas où q est pair (resp. impair), nous avons

$$f(x; \theta) = f(x; \theta_0), \quad \forall x \in \mathbb{Z}.$$

- Supposons que $s(\alpha_0 x) = -1$ et $s(\lambda_0) = -1$. Rappelons que pour tout $a \in \mathbb{R}$, nous avons $\langle -a \rangle = -\langle a \rangle$. Ainsi, pour tout $x \in \mathbb{Z}$, nous obtenons

$$f(x; \theta) - f(x; \theta_0) = -(\langle \{ \alpha_0 x \} + \{ \lambda \} \rangle - \langle \{ \alpha_0 x \} + \{ \lambda_0 \} \rangle) = 0.$$

- Supposons que $s(\alpha_0 x) = 1$ et $s(\lambda_0) = -1$. Nous avons,

$$f(x; \theta) - f(x; \theta_0) = \langle \{ \alpha_0 x \} - \{ \lambda \} \rangle - \langle \{ \alpha_0 x \} - \{ \lambda_0 \} \rangle.$$

Ainsi, d'après (2.63) (resp. (2.64)) dans le cas où q est pair (resp. impair), nous obtenons

$$f(x; \theta) = f(x; \theta_0), \forall x \in \mathbb{Z}.$$

- Supposons que $s(\alpha_0 x) = -1$ et $s(\lambda_0) = 1$. Nous avons,

$$f(x; \theta) - f(x; \theta_0) = -(\langle \{\alpha_0 x\} - \{\lambda\} \rangle - \langle \{\alpha_0 x\} - \{\lambda_0\} \rangle).$$

Ainsi, nous trouvons

$$f(x; \theta) = f(x; \theta_0), \forall x \in \mathbb{Z}.$$

Maintenant, notre objectif est de montrer que pour tout $\theta = (\alpha, \lambda) \in \Theta$,

$$\langle \alpha x + \lambda \rangle = \langle \alpha_0 x + \lambda_0 \rangle, \forall x \in \mathbb{Z} \Rightarrow \theta \in A_0$$

Soit $\theta \notin A_0$. Nous distinguons trois cas :

- Supposons que $\alpha \neq \alpha_0$. Pour ce cas nous nous référons à la preuve de la Proposition 2.4.2. Ainsi, alors il existe $x_0 \in \mathbb{Z}$ tel que $|f(x_0; \theta) - f(x_0; \theta_0)| > 0$.
- Supposons que $\alpha = \alpha_0$ et $\langle \lambda \rangle \neq \langle \lambda_0 \rangle$. Il existe $x_0 \in \mathbb{Z}$ tel que $\alpha_0 x_0 = \frac{p}{q} x_0 \in \mathbb{Z}$.

Par conséquent, nous obtenons $\{\alpha_0 x_0\} = 0$. Ainsi,

$$|f(x_0; \theta) - f(x_0; \theta_0)| = |\langle \lambda \rangle - \langle \lambda_0 \rangle| \geq 1.$$

- Supposons que $\alpha = \alpha_0$, $\langle \lambda \rangle = \langle \lambda_0 \rangle$ et $\lambda \notin I_0$. Pour simplifier les notations, nous supposons sans perte de généralité $\alpha_0 > 0$ et $\lambda \geq \lambda_0 \geq 0$.

Premièrement, nous considérons le cas où q est pair. Rappelons qu'il existe

$k_0 \in \{0, \dots, q-1\}$ tel que

$$i_0 = \left[\frac{k_0}{q}, \frac{k_0 + 1}{q} \right[.$$

Nous distinguons trois cas :

- Si $i_0 \subset \left[0, \frac{1}{2} \right[$. Comme $\lambda \geq \lambda_0$ et $\{\lambda\} \notin i_0$, il est clair qu'ici $q \geq 4$ et $k_0 \neq \frac{q}{2} - 1$. Ainsi, $\{\lambda\} \in \left[0, \frac{1}{2} \right[$ et $\{\lambda\} > \{\lambda_0\}$ ($\{\lambda\} > i_0$). Donc, il existe $x_0 \in \left\{ 1, \dots, \frac{q}{2} - 1 \right\}$ tel que

$$f(x_0; \theta) - f(x_0; \theta_0) = \langle \{\alpha_0 x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle = 1.$$

Par exemple, pour $q = 4$, nous avons $i_0 = \left[0, \frac{1}{4} \right[$ et par suite $\{\lambda\} \in \left[\frac{1}{4}, \frac{1}{2} \right[$.

Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{4}$ et par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 \subset \left[\frac{1}{2}, 1\right[$ et $\{\lambda\} \in \left[\frac{1}{2}, 1\right[$. Comme $\lambda \geq \lambda_0$ et $\{\lambda\} \notin i_0$, il est clair qu'ici $q \geq 4$ et $k_0 \neq q - 1$. Ainsi, nous avons $\{\lambda\} > \{\lambda_0\}$ ($\{\lambda\} > i_0$) et par conséquent il existe $x_0 \in \left\{\frac{q}{2} + 1, \dots, q - 1\right\}$ tel que

$$f(x_0; \theta) - f(x_0; \theta_0) = \langle \{\alpha_0 x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle = 1.$$

Par exemple, pour $q = 4$, nous avons $i_0 = \left[\frac{1}{2}, \frac{3}{4}\right[$ et par suite $\{\lambda\} \in \left[\frac{3}{4}, 1\right[$.

Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{3}{4}$ et par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 \subset \left[\frac{1}{2}, 1\right[$ et $\{\lambda\} \in \left[0, \frac{1}{2}\right[$. Ainsi, $[\lambda] = [\lambda_0] + 1$. Donc, pour tout $x \in \mathbb{N}$, nous avons

$$f(x; \theta) - f(x; \theta_0) = \langle \{\alpha_0 x\} + \{\lambda\} \rangle - \langle \{\alpha_0 x\} + \{\lambda_0\} - 1 \rangle.$$

Comme q est pair, alors il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{2}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

Notons que $q = 2$ est un cas particulier. Si $\{\lambda_0\} \in i_0 = \left[0, \frac{1}{2}\right[$, alors sous les hypothèses $\lambda \geq \lambda_0 \geq 0$ et $\langle \lambda \rangle = \langle \lambda_0 \rangle$, nous avons

$$\{\lambda\} \in i_0.$$

Si $\{\lambda_0\} \in i_0 = \left[\frac{1}{2}, 1\right[$, alors

$$\{\lambda\} \in \left[0, \frac{1}{2}\right[\text{ et } [\lambda] = [\lambda_0] + 1.$$

Par suite, il existe $x_0 \in \mathbb{Z}$ tel que $\{\alpha_0 x_0\} = \frac{1}{2}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

Maintenant, nous considérons le cas où q est impair et $\{\lambda_0\} \in \left[\frac{1}{2q}, \frac{2q-1}{2q}\right[$.

Rappelons qu'il existe $k_0 \in \{1, 3, \dots, q, \dots, 2q-3\}$ tel que

$$i_0 = \left[\frac{k_0}{2q}, \frac{k_0+2}{2q}\right[.$$

Pour simplifier les notations, nous posons $q = 5$. Rappelons que $\lambda \geq \lambda_0$, $\langle \lambda \rangle = \langle \lambda_0 \rangle$ et $\{\lambda\} \notin I_0$. Ainsi, nous distinguons les cas suivants :

- Si $i_0 \subset \left[0, \frac{1}{2}\right[$. Comme $\lambda \geq \lambda_0$, alors

$$i_0 = \left[\frac{1}{10}, \frac{3}{10}\right[\text{ et } \{\lambda\} \in \left[\frac{3}{10}, \frac{1}{2}\right[$$

Par suite, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{1}{2}, \frac{7}{10}\right[$ et $\{\lambda\} \in \left[\frac{7}{10}, \frac{9}{10}\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{4}{5}$.

Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{1}{2}, \frac{7}{10}\right[$ et $\{\lambda\} \in \left[\frac{9}{10}, 1\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{3}{5}$.

Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{7}{10}, \frac{9}{10}\right[$ et $\{\lambda\} \in \left[\frac{9}{10}, 1\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{3}{5}$.

Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{1}{2}, \frac{7}{10}\right[$ (resp. $\left[\frac{7}{10}, \frac{9}{10}\right[$) et $\{\lambda\} \in \left[\frac{3}{10}, \frac{1}{2}\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{1}{2}, \frac{7}{10}\right[$ (resp. $\left[\frac{7}{10}, \frac{9}{10}\right[$) et $\{\lambda\} \in \left[\frac{1}{10}, \frac{3}{10}\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{2}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $i_0 = \left[\frac{1}{2}, \frac{7}{10}\right[$ (resp. $\left[\frac{7}{10}, \frac{9}{10}\right[$) et $\{\lambda\} \in \left[0, \frac{1}{10}\right[$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{3}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

Maintenant, supposons que $\{\lambda_0\} \in \left[\frac{9}{10}, 1\right[$. Rappelons que $\lambda \geq \lambda_0 \geq 0$ et

$\lambda \notin I_0 = \left[\lambda_0\right] + \frac{2q-1}{2q}, [\lambda_0] + 1 + \frac{1}{2q}\right[$. Ainsi, nous notons deux cas possible :

- Si $\{\lambda\} \in \left[\frac{3}{10}, \frac{1}{2}\right]$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $\{\lambda\} \in \left[\frac{1}{10}, \frac{3}{10}\right]$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{2}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

Finalement, supposons que $\{\lambda_0\} \in \left[0, \frac{1}{10}\right]$ et $[\lambda_0] \geq 1$ (resp. $[\lambda_0] = 0$). Par suite, nous avons

$$I_0 = \left[[\lambda_0] - 1 + \frac{2q-1}{2q}, [\lambda_0] + \frac{1}{2q}\right] \text{ (resp. } I_0 = \left[-\frac{1}{10}, \frac{1}{10}\right]). \text{ Ainsi,}$$

- Si $\{\lambda\} \in \left[\frac{3}{10}, \frac{1}{2}\right]$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{1}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

- Si $\{\lambda\} \in \left[\frac{1}{10}, \frac{3}{10}\right]$. Ainsi, il existe $x_0 \in \mathbb{N}$ tel que $\{\alpha_0 x_0\} = \frac{2}{5}$. Par conséquent,

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

□

Lemme 2.4.3. Soit $x_0 \in \mathbb{N}^*$ fixé, tel que $x_0 \neq kq$ pour tout $k \in \mathbb{N}$. la fonction $\alpha \longrightarrow \{\alpha x_0\}$, définie sur \mathbb{R} et à valeurs dans $[0, 1[$, est continue en $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$ et $p \wedge q = 1$.

Preuve. Soit x_0 fixé. Nous avons, $\{\alpha x_0\} = |\alpha x_0| - [\alpha x_0]$, où $[\cdot]$ est la fonction partie entière. Alors la fonction $\{\cdot\} : \mathbb{R} \longrightarrow [0, 1[$, définie par $\alpha \longrightarrow \{\alpha x_0\}$, est discontinue en α si $\alpha x_0 \in \mathbb{Z}$. Par conséquent, si $\alpha_0 = \frac{p}{q}$ et $x_0 \neq kq$ pour tout $k \in \mathbb{N}$, alors $\alpha_0 x_0 \notin \mathbb{Z}$.

Ainsi, $\alpha \longrightarrow \{\alpha x_0\}$ est continue en α_0 .

□

Rappelons que sous [H1] et d'après la Proposition 2.4.1 nous avons,

$$K(\theta) - K(\theta_0) = \sigma_{\theta_0} [g_{\theta}(\cdot) - g_{\theta_0}(\cdot)] = \mu_{\theta_0} [f(\cdot; \theta) - f(\cdot; \theta_0)]^2,$$

où

$$g_{\theta}(y) = [z - f(x; \theta)]^2, \quad \forall y = (x, z) \in E^2$$

et

$$f(x; \theta) = \langle \alpha x + \lambda \rangle, \quad \forall x \in E \text{ et } \theta = (\alpha, \lambda) \in \Theta.$$

De plus, d est la distance sur Θ définie par

$$d(\theta, \theta') = \max \left\{ |\alpha - \alpha'|, |\lambda - \lambda'| \right\}, \quad \forall \theta = (\alpha, \lambda), \theta' = (\alpha', \lambda') \in \Theta.$$

Nous posons,

$$A_0 = \{\theta, \alpha = \alpha_0 \text{ et } \lambda \in I_0\}. \quad (2.71)$$

Rappelons que $\theta_0 \in A_0$. Soit $\theta_* \in A_0$. D'après la Proposition 2.4.4, nous avons

$$K(\theta_*) - K(\theta_0) = \mu_{\theta_0} [f(\cdot; \theta_*) - f(\cdot; \theta_0)]^2 = 0. \quad (2.72)$$

Proposition 2.4.5. *Supposons que [H1] soit vérifiée. Si $\alpha_0 = \frac{p}{q}$, où $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$ et $p \wedge q = 1$, alors pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_\varepsilon^0} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon^0 = \{\theta : d(\theta, A_0) \geq \varepsilon\}$.

Preuve. Pour simplifier les notations, nous considérons le cas où q est pair.

Donc, il existe $k_0 \in \{0, 1, \dots, q-1\}$ tel que

$$\{\lambda_0\} \in i_0 = \left[\frac{k_0}{q}, \frac{k_0+1}{q} \right[.$$

Par suite,

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda_0 \rangle \text{ et } \{\lambda\} \in i_0\} \text{ et } A_0 = \{\theta, \alpha = \alpha_0 \text{ et } \lambda \in I_0\}.$$

Nous distinguons trois cas pour l'événement Θ_ε^0 . Nous posons, $\Theta_\varepsilon^0 = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, où

$$\Gamma_1 = \{\theta, |\alpha - \alpha_0| \geq \varepsilon\}, \quad (2.73)$$

$$\Gamma_2 = \{\theta, |\alpha - \alpha_0| < \varepsilon, d(\lambda, I_0) \geq \varepsilon, \langle \lambda \rangle \neq \langle \lambda_0 \rangle\}, \quad (2.74)$$

et

$$\Gamma_3 = \{\theta, |\alpha - \alpha_0| < \varepsilon, d(\lambda, I_0) \geq \varepsilon, \langle \lambda \rangle = \langle \lambda_0 \rangle\}. \quad (2.75)$$

Notons que,

$$d(\lambda, I_0) = \inf_{\lambda_* \in I_0} |\lambda - \lambda_*|.$$

L'idée de la preuve est basée sur l'équation (2.32). Alors, le but est de trouver $x_0 \in E$ tel que

$$|f(x_0; \theta) - f(x_0; \theta_0)| > 0, \text{ uniformément sur } \Gamma_i, \quad i = 1, 2, 3.$$

1. Nous considérons le premier cas, $\theta \in \Gamma_1$. Comme Θ est compact et le support de μ_{θ_0} n'est pas borné, alors il existe $x_0 > 0$ tel que $\forall |x| \geq x_0$, nous avons

$$|f(x; \theta) - f(x; \theta_0)| \geq \frac{\varepsilon}{2}|x|, \text{ uniformément sur } \Gamma_1. \text{ Ainsi,}$$

$$\inf_{\theta \in \Gamma_1} |K(\theta) - K(\theta_0)| \geq \int_{|x| \geq x_0} \left(\frac{\varepsilon}{2}x\right)^2 \mu_{\theta_0}(dx) > 0.$$

2. Supposons que $\theta \in \Gamma_2$. Pour $x_0 = 0$, nous avons

$$|f(0; \theta) - f(0; \theta_0)| = |\langle \lambda \rangle - \langle \lambda_0 \rangle| \geq 1.$$

Ainsi,

$$\inf_{\theta \in \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

3. Nous considérons le cas où $\theta \in \Gamma_3$. Nous supposons sans perte de généralité que $\alpha_0 = \frac{1}{q} > 0$ et $\lambda_0 \geq 0$. Nous avons,

$$\{\lambda_0\} \in i_0 = \left[\frac{k_0}{q}, \frac{k_0 + 1}{q} \right[\text{ et par suite } I_0 = \left[[\lambda_0] + \frac{k_0}{q}, [\lambda_0] + \frac{k_0 + 1}{q} \right[.$$

Nous distinguons quatre cas dépendants de la position de la partie fractionnaire de λ_0 et celle de λ :

- (a) Cas A : $i_0 \subset [0, \frac{1}{2}[$ et $\lambda \geq \lambda_0$.
- (b) Cas B : $i_0 \subset [\frac{1}{2}, 1[$ et $\lambda \geq \lambda_0$.
- (c) Cas C : $i_0 \subset [0, \frac{1}{2}[$ et $\lambda \leq \lambda_0$.
- (d) Cas D : $i_0 \subset [\frac{1}{2}, 1[$ et $\lambda \leq \lambda_0$.

– Cas A : $i_0 \subset [0, \frac{1}{2}[$, $\lambda \geq \lambda_0$. Il est simple de vérifier qu'ici

$$q \geq 4, \{\lambda\} \in \left[0, \frac{1}{2} \right[, k_0 \in \left\{ 0, 1, \dots, \frac{q}{2} - 2 \right\} \text{ et } \{\lambda\} \geq \frac{k_0 + 1}{q}.$$

Par suite,

$$d(\theta, A_0) = d(\lambda, I_0) = \inf_{\{\lambda_*\} \in i_0} \{\lambda\} - \{\lambda_*\} \geq \varepsilon. \quad (2.76)$$

Ainsi, il existe $l \in \left\{ k_0 + 1, \dots, \frac{q}{2} - 1 \right\}$ tel que $\{\lambda\} \in \left[\frac{l}{q}, \frac{l+1}{q} \right[$.

Soit $x_0 = \frac{q}{2} - l$ fixé. Alors, pour tout $\{\lambda_*\} \in i_0$ nous obtenons

$$\{\alpha_0 x_0\} + \{\lambda_*\} < \frac{1}{2}, \quad (2.77)$$

et d'après (2.76) nous trouvons

$$\{\alpha_0 x_0\} + \{\lambda\} \geq \frac{1}{2} + \varepsilon. \quad (2.78)$$

Comme $x_0 = \frac{q}{2} - l \neq kq$ où $k \in \mathbb{N}$, d'après Lemme 2.4.3, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue en $\alpha_0 = \frac{p}{q}$. Ainsi, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$ nous trouvons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0]. \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \varepsilon. \end{cases} \quad (2.79)$$

D'après (2.79) et (2.78) nous trouvons

$$\begin{aligned}\{\alpha x_0\} + \{\lambda\} &= (\{\alpha x_0\} - \{\alpha x_0\}) + (\{\alpha x_0\} + \{\lambda\}) \\ &\geq -\varepsilon + \left(\frac{1}{2} + \varepsilon\right) = \frac{1}{2}.\end{aligned}$$

Alors, pour tout $\theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$, nous avons

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

– Case B : $i_0 \subset [\frac{1}{2}, 1[$ et $\lambda \geq \lambda_0$. Nous distinguons deux sous-cas :

– Supposons que $\{\lambda\} \in [\frac{1}{2}, 1[$. Ici, il est simple de vérifier que

$$q \geq 4, \quad k_0 \in \left\{\frac{q}{2}, \dots, q-2\right\}, \text{ et } \{\lambda\} \geq \frac{k_0 + 1}{q}.$$

Par conséquent,

$$d(\theta, A_0) = d(\lambda, I_0) = \inf_{\{\lambda_*\} \in i_0} \{\lambda\} - \{\lambda_*\} \geq \varepsilon. \quad (2.80)$$

Ainsi, il existe $x_0 \in \{\frac{q}{2} + 1, \dots, q-1\}$ tel que pour tout $\{\lambda_*\} \in i_0$ nous avons

$$\{\alpha_0 x_0\} + \{\lambda_*\} < \frac{3}{2}, \quad (2.81)$$

et d'après (2.80) nous trouvons

$$\{\alpha_0 x_0\} + \{\lambda\} \geq \frac{3}{2} + \varepsilon. \quad (2.82)$$

Comme $x_0 \neq kq$ où $k \in \mathbb{N}$, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue $\alpha_0 = \frac{p}{q}$

(voir Lemme 2.4.3). Alors, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$ nous obtenons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0]. \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \varepsilon. \end{cases} \quad (2.83)$$

d'après (2.83) et (2.82) nous trouvons

$$\begin{aligned}\{\alpha x_0\} + \{\lambda\} &= (\{\alpha x_0\} - \{\alpha_0 x_0\}) + (\{\alpha_0 x_0\} + \{\lambda\}) \\ &\geq -\varepsilon + \left(\frac{3}{2} + \varepsilon\right) = \frac{3}{2}.\end{aligned}$$

Alors, pour tout $\theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta(\varepsilon, \alpha_0, \lambda_0)\}$ nous avons

$$f(x_0; \theta) - f(x_0; \theta_0) = 1.$$

– Supposons que $\{\lambda\} \in \left[0, \frac{1}{2}\right]$. Ici, il est simple de vérifier que

$$q \geq 2, \text{ et } k_0 \in \left\{\frac{q}{2}, \dots, q-1\right\}.$$

Par conséquent, nous avons

$$d(\lambda, I_0) = \inf_{\{\lambda_*\} \in i_0} \{\lambda\} - \{\lambda_*\} + 1 \geq \varepsilon \text{ et } [\lambda] = [\lambda_*] + 1. \quad (2.84)$$

Ainsi, pour $x_0 = \frac{q}{2}$ et pour tout $\{\lambda_*\} \in i_0$ nous avons

$$1 \leq \{\alpha_0 x_0\} + \{\lambda_*\} < \frac{3}{2}. \quad (2.85)$$

De plus, d'après (2.84), nous trouvons

$$\{\alpha_0 x_0\} + \{\lambda\} \geq \frac{1}{2} + \varepsilon. \quad (2.86)$$

Pour ce x_0 fixé, la fonction $\alpha \rightarrow \{\alpha x_0\}$ est continue en $\alpha_0 = \frac{p}{q}$

(voir Lemme 2.4.3). Alors, il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ tel que $\forall |\alpha - \alpha_0| \leq \eta$ nous obtenons

$$\begin{cases} [\alpha x_0] = [\alpha_0 x_0]. \\ |\{\alpha x_0\} - \{\alpha_0 x_0\}| \leq \varepsilon. \end{cases} \quad (2.87)$$

D'après (2.87) et (2.86) nous trouvons

$$\begin{aligned} \{\alpha x_0\} + \{\lambda\} &= (\{\alpha x_0\} - \{\alpha_0 x_0\}) + (\{\alpha_0 x_0\} + \{\lambda\}) \\ &\geq -\varepsilon + \left(\frac{1}{2} + \varepsilon\right) = \frac{1}{2}. \end{aligned}$$

Ainsi, pour tout $\theta \in \Gamma'_3 = \{\theta \in \Gamma_3, |\alpha - \alpha_0| \leq \eta\}$ nous trouvons

$$\begin{aligned} f(x_0; \theta) - f(x_0; \theta_0) &= \langle \alpha x_0 + \lambda \rangle - \langle \alpha_0 x_0 + \lambda_0 \rangle, \\ &= \langle [\alpha x_0] + [\lambda] + \{\alpha x_0\} + \{\lambda\} \rangle \\ &\quad - \langle [\alpha_0 x_0] + [\lambda_0] + \{\alpha_0 x_0\} + \{\lambda_0\} \rangle, \\ &= ([\alpha x_0] - [\alpha_0 x_0]) + ([\lambda] - [\lambda_0]) \\ &\quad + (\langle \{\alpha x_0\} + \{\lambda\} \rangle - \langle \{\alpha_0 x_0\} + \{\lambda_0\} \rangle) \\ &= 1. \end{aligned}$$

Les cas C et D peuvent être traités de façon similaire, avec un bon choix de x_0 et α suffisamment proche de α_0 . Alors, nous déduisons qu'il existe $\eta = \eta(\varepsilon, \alpha_0, \lambda_0) \leq \varepsilon$ et une constante $e_0 > 0$ tels que uniformément sur Γ'_3 nous avons

$$|K(\theta) - K(\theta_0)| \geq e_0 > 0.$$

Pour $\theta \in \Gamma_3'' = \Gamma_3 \setminus \Gamma_3' = \{\theta \in \Gamma_3, |\alpha - \alpha_0| > \eta, \langle \lambda \rangle = \langle \lambda_0 \rangle\}$ les mêmes arguments pour Γ_1 peuvent être appliqués ici et par suite il existe $d_0 > 0$ tel que

$$|K(\theta) - K(\theta_0)| \geq d_0 > 0.$$

□

Théorème 2.4.2. *Supposons que*

1. α_0 , la vraie valeur de α , est un nombre rationnel appartenant l'intervalle $] -1, 1[$, où $\alpha_0 = \frac{p}{q}$, avec $p \in \mathbb{Z}$ $q \in \mathbb{N}^*$ et $p \wedge q = 1$;
2. Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible ;
3. Pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;
4. L'espace des paramètres Θ , est un compact sous-ensemble $] -1, 1[\times \mathbb{R}$.

Alors,

$$d(\hat{\theta}_n, A_0) \xrightarrow{p.s.} 0, \text{ où } A_0 = \{\theta \in \Theta, \alpha = \alpha_0 \text{ et } \lambda \in I_0\}.$$

En d'autres termes, $\hat{\alpha}_n$ est fortement consistant tandis que $\hat{\lambda}_n$ converge vers un intervalle de taille $\frac{1}{q}$ contenant λ_0 .

Preuve. Rappelons que, d'après la Proposition 2.4.1 et 2.4.5, nous avons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0$$

et

$$\inf_{\theta \in \Theta_\varepsilon^0} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon^0 = \{\theta : d(\theta, A_0) \geq \varepsilon\}$. La conclusion $\hat{\theta}_n \rightarrow \theta_0$ presque sûrement résulte par des arguments standards de la théorie des M-estimateurs (voir par exemple, Van Der Vaart (1998, [61]), Théorème 5.7). □

2.5 Méthode numérique pour calculer $\hat{\theta}_n$

Soient X_0, X_1, \dots, X_n des observations du processus RINAR(1). Notre objectif est de trouver une méthode numérique pour calculer l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \varphi_n(\theta),$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n [X_t - f(X_{t-1}; \theta)]^2 = \frac{1}{n} \sum_{t=1}^n [X_t - \langle \alpha X_{t-1} + \lambda \rangle]^2.$$

La principale difficulté pour le calcul de $\hat{\theta}_n$ provient de la discontinuité de la fonction de contraste. Les méthodes classiques, comme la descente du gradient, ne sont pas valables. Dans cette section, nous proposons un algorithme basé sur la recherche dichotomique successive.

En premier lieu, nous expliquons l'idée de cet algorithme dans un cadre simple.

Soit h une fonction réelle irrégulière définie sur \mathbb{R} , appelée la fonction objective.

Notre but est de trouver

$$\hat{\theta}_n = \arg \min_{\theta \in [a, b]} h(\theta).$$

Soit θ_0 notre point de départ. Ainsi, l'intervalle initial de recherche est $[a, b]$.

Ainsi, nous notons $a_0 = a$, $b_0 = b$ et $N_0 = b_0 - a_0$. Par suite, le passage

de l'étape i à $i + 1$, pour $i = 0, 1, \dots$, se fait de la manière suivante :

Soient $c_i = \frac{\theta_i + a_i}{2}$ et $d_i = \frac{\theta_i + b_i}{2}$. Nous calculons $h(c_i)$, $h(\theta_i)$ et $h(d_i)$.

Ensuite, selon le minimum de ces derniers, nous distinguons trois cas (voir Figure 2.7) :

- Si $h(\theta_i) < \min\{h(c_i), h(d_i)\}$, alors $a_{i+1} = c_i$, $\theta_{i+1} = \theta_i$ et $b_{i+1} = d_i$.
- Si $h(c_i) < \min\{h(\theta_i), h(d_i)\}$, alors $a_{i+1} = a_i$, $\theta_{i+1} = c_i$ et $b_{i+1} = \theta_i$.
- Si $h(d_i) < \min\{h(\theta_i), h(c_i)\}$, alors $a_{i+1} = \theta_i$, $\theta_{i+1} = d_i$ et $b_{i+1} = b_i$.

Le processus s'arrête à l'étape $i + 1$ si $N_{i+1} = b_{i+1} - a_{i+1} \leq 0.001$.

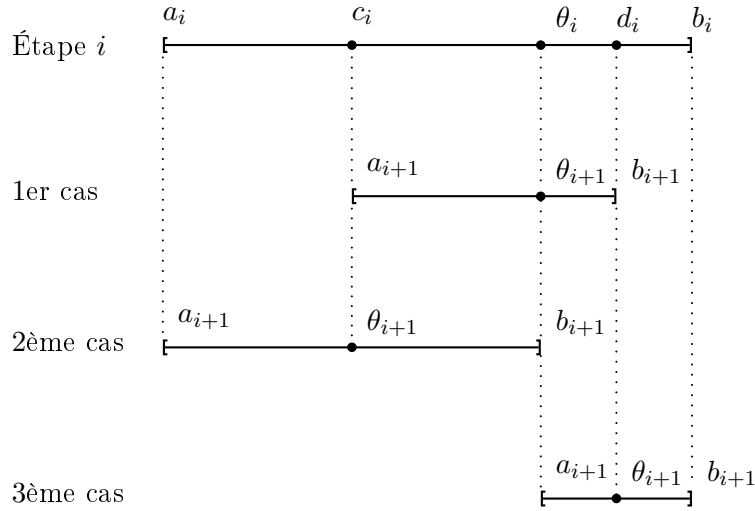


FIG. 2.7 – Le passage de l'étape i à $i + 1$ pour la recherche dichotomique.

2.5.1 Initialisation

Comme le processus RINAR(1) est inspiré du modèle AR(1) réel, nous proposons pour le point de départ de l'algorithme $\hat{\theta}_0 = (\hat{\alpha}_0, \hat{\lambda}_0)$, l'estimateur de Yule-Walker du paramètre $\theta = (\alpha, \lambda)$. Ainsi, nous avons

$$\hat{\alpha}_0 = \hat{\rho}(1) \text{ et } \hat{\lambda}_0 = \bar{X}_n (1 - \hat{\alpha}_0).$$

Ici, $\hat{\rho}(1)$ est le coefficient d'autocorrélation empirique du premier ordre et \bar{X}_n est la moyenne empirique.

2.5.2 Recherche dichotomique successive

Comme le paramètre θ possède deux composants, la transition de $\hat{\theta}_k = (\hat{\alpha}_k, \hat{\lambda}_k)$ à $\hat{\theta}_{k+1} = (\hat{\alpha}_{k+1}, \hat{\lambda}_{k+1})$, pour $k = 0, 1, \dots$, se fait en deux phases. Ainsi, nous appliquons une recherche dichotomique successive sur α puis sur λ .

Dans la première phase nous fixons $\hat{\lambda}_k$. Par suite, une recherche dichotomique sur α nous conduit à la nouvelle valeur $\hat{\alpha}_{k+1}$. Ici, nous utilisons $] -1, 1[$ comme intervalle initial

de la recherche. Ainsi, nous notons $left = -1$, $right = 1$ et $range = |right - left|$. Soient $mid-left = \frac{left + \hat{\alpha}_k}{2}$ et $mid-right = \frac{\hat{\alpha}_k + right}{2}$, où $mid-left$ et $mid-right$ représentent respectivement les milieux à gauche et à droite de $\hat{\alpha}_k$.

Puis, nous calculons les valeurs suivantes :

$$\varphi_n \left(mid-left, \hat{\lambda}_k \right), \varphi_n \left(\hat{\alpha}_k, \hat{\lambda}_k \right) \text{ et } \varphi_n \left(mid-right, \hat{\lambda}_k \right).$$

Selon le minimum de ces dernières, $left$, $\hat{\alpha}_k$, et $right$ changent leurs valeurs actuelles.

Par exemple, Si $\varphi_n \left(\hat{\alpha}_k, \hat{\lambda}_k \right)$ est le minimum, alors $left$ prend la valeur de $mid-left$ et $right$ prend celle de $mid-right$. Par suite, le processus s'arrête quand $range \leq 0.001$.

Le pseudo-code suivant définit la recherche dichotomique utilisée pour aboutir à $\hat{\alpha}_{k+1}$.

```

left <- -1;
right <- 1;
range <- 2;
mid_left <- ( left + previous_alpha ) / 2;
mid_right <- ( right + previous_alpha ) / 2;
while range > 0.001 do
  begin
    V_1 <- Phi_n ( alpha_k , lambda_k );
    V_2 <- Phi_n ( mid_left , lambda_k );
    V_3 <- Phi_n ( mid_right , lambda_k );
    i <- j such that V_j is min ( V_1, V_2, V_3 )
    case i of
      1 : begin left <- mid_left; right <- mid_right end;
      2 : begin right <- alpha_k; alpha_k <- mid_left end;
      3 : begin left <- alpha_k; alpha_k <- mid_right end;
    range <- abs ( right - left )
  end;
end;
```

Dans la deuxième phase, nous fixons $\hat{\alpha}_{k+1}$. Par suite, une recherche dichotomique sur λ nous conduit à la nouvelle valeur $\hat{\lambda}_{k+1}$. Ici nous proposons $\left[\hat{\lambda}_0 - 5 |\hat{\lambda}_0|, \hat{\lambda}_0 + 5 |\hat{\lambda}_0| \right]$ comme intervalle initial de la recherche, lequel nous semble assez large pour couvrir la plupart des situations.

Ainsi, nous notons $left = \hat{\lambda}_0 - 5 |\hat{\lambda}_0|$, $right = \hat{\lambda}_0 + 5 |\hat{\lambda}_0|$ et $range = |right - left|$.

Soient $mid-left = \frac{left + \hat{\lambda}_k}{2}$ et $mid-right = \frac{\hat{\lambda}_k + right}{2}$, où $mid-left$ et $mid-right$ représentent respectivement les milieux à gauche et à droite de $\hat{\lambda}_k$.

Puis, nous calculons les valeurs suivantes :

$$\varphi_n(\hat{\alpha}_{k+1}, mid-left), \varphi_n(\hat{\alpha}_{k+1}, \hat{\lambda}_k) \text{ et } \varphi_n(\hat{\alpha}_{k+1}, mid-right).$$

Selon le minimum de ces dernières, *left*, $\hat{\lambda}_k$, et *right* changent leurs valeurs actuelles.

Enfin, le processus s'arrête quand $range \leq 0.001$. Le pseudo-code suivant définit la recherche dichotomique utilisée pour aboutir à $\hat{\lambda}_{k+1}$.

```

left <- lambda_0 - 5 |lambda_0|;
right <- lambda_0 + 5 |lambda_0|;
range <- 10|lambda_0|;
mid_left <- ( left + previous_lambda ) / 2;
mid_right <- ( right + previous_lambda ) / 2;
while range > 0.001 do
  begin
    V_1 <- Phi_n ( alpha_k+1 , lambda_k );
    V_2 <- Phi_n ( alpha_k+1 , mid_left );
    V_3 <- Phi_n ( alpha_k+1 , mid_right );
    i <- j such that V_j is min ( V_1, V_2, V_3 )
    case i of
      1 : begin left <- mid_left; right <- mid_right end;
      2 : begin right <- lambda_k; lambda_k <- mid_left end;
      3 : begin left <- lambda_k; lambda_k <- mid_right end;
    range <- abs ( right - left )
  end;

```

Ainsi, nous atteignons la $(k+1)$ ^{ème} itération.

La recherche s'arrête quand les résultats de deux itérations consécutives sont très proches. Plus précisément, nous nous arrêtons à la k ^{ème} itération si

$$d\left((\hat{\alpha}_k, \hat{\lambda}_k), (\hat{\alpha}_{k+1}, \hat{\lambda}_{k+1})\right) = \max\left\{|\hat{\alpha}_{k+1} - \hat{\alpha}_k|, |\hat{\lambda}_{k+1} - \hat{\lambda}_k|\right\} \leq 0.001. \quad (2.88)$$

Notons que, dans le cas général, ce critère d'arrêt est satisfait après peu d'itérations.

À la fin des itérations, nous obtenons le couple $(\hat{\alpha}_n, \hat{\lambda}_n)$ qui minimise la fonction objectif (de contraste) φ_n .

Nous tenons à mentionner que l'estimation de Yule-Walker $\hat{\theta}_0 = (\hat{\alpha}_k, \hat{\lambda}_k)$ qui représente le point de départ de notre recherche est en général une très bonne proposition initiale de notre algorithme de minimisation.

Pour cela, nous pensons que l'intervalle initial $\hat{\lambda}_0 \pm 5|\hat{\lambda}_0|$ proposé pour la recherche des

$\hat{\lambda}_k$ est plus que suffisant. Clairement, sans une si bonne estimation initiale, nous pouvons employer une autre stratégie plus sophistiquée que celle proposée pour définir l'intervalle initial de la recherche à chaque itération, voir par exemple Press & al. (2007, [57]).

2.5.3 Étude de simulation

Notre objectif est d'illustrer les performances de la méthode choisie pour calculer $\hat{\theta}_n$. Pour cela, nous simulons 500 observations du processus RINAR(1), avec $\theta_0 = (\alpha_0, \lambda_0)$ fixé. Pour obtenir un bruit centré à valeurs dans \mathbb{Z} , nous considérons deux suites ξ_t et η_t indépendantes de variables i.i.d. suivant une loi de Poisson de paramètre μ , et nous posons $\varepsilon_t = \xi_t - \eta_t$. Puis, en utilisant l'algorithme indiqué précédemment, nous calculons $\hat{\theta}_n$. Après 500 répétitions indépendantes, nous calculons la moyenne et l'écart-type des suites des estimations obtenues $\{\hat{\alpha}_{n,i}, 1 \leq i \leq 500\}$ et $\{\hat{\lambda}_{n,i}, 1 \leq i \leq 500\}$, et dessinons leurs histogrammes.

Ici nous proposons trois jeux de paramètres correspondant au choix de trois différentes valeurs de θ_0 . Notons qu'inévitablement, les α_0 utilisés pour les simulations sont rationnels. Par conséquent, comme prévu par le Théorème 2.4.2, la consistance de l'estimateur des moindres carrés $\hat{\theta}_n$ doit être jugée en fonction d'une fluctuation inévitable autour de λ_0 .

– Cas A : $\alpha_0 = -0.4$ et $\lambda_0 = 1.44$.

Donc, $\alpha_0 = \frac{p}{q}$ où $p = -2$ et $q = 5$ est impair. Ainsi, nous avons $\{\lambda_0\} = 0.44 \in i_0 = [\frac{3}{10}, \frac{5}{10}[$ et par suite $I_0 = [1.3, 1.5[$. La moyenne des $\hat{\alpha}_{n,i}$ est -0.4075 et leur écart-type est 0.0454 . D'ailleurs, la moyenne des $\hat{\lambda}_{n,i}$ est 1.3988 et leur écart-type est 0.1626 . Notons que la moyenne des estimations de λ appartient à l'intervalle I_0 .

Les figures suivantes nous donnent les histogrammes des estimations obtenues.

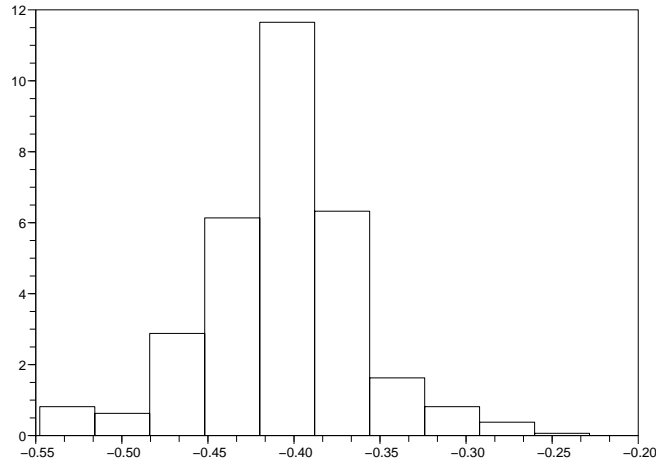


FIG. 2.8 – L'histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = -0.4$.

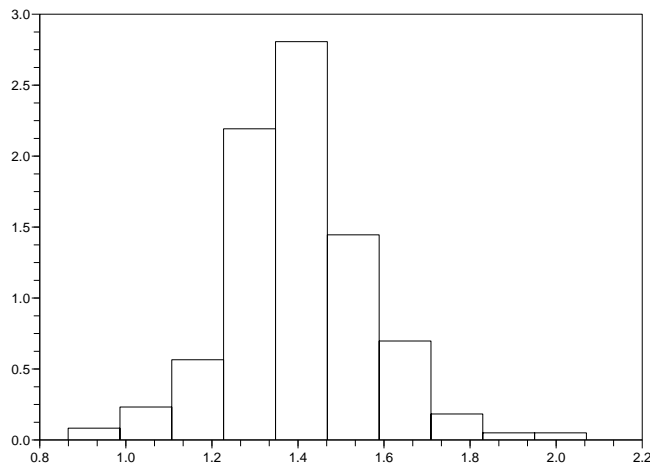


FIG. 2.9 – L'histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = 1.44$.

– Cas B : $\alpha_0 = -0.2$ et $\lambda_0 = -2.96$.

Donc, $\alpha_0 = \frac{p}{q}$ où $p = -1$ et $q = 5$ est impair. Ainsi, nous avons $\{\lambda_0\} = 0.96 \in i_0 = [\frac{9}{10}, 1[$ et par suite $I_0 = [-2.9, -3.1[$. La moyenne des $\hat{\alpha}_{n,i}$ est -0.2038 et leur écart-type est 0.0545 . D'autre part, la moyenne des $\hat{\lambda}_{n,i}$ est -2.9953 et leur écart-type est 0.1861 . Notons que la moyenne des estimations de λ appartient à l'intervalle I_0 .

Les figures suivantes nous donnent les histogrammes des estimations obtenues.

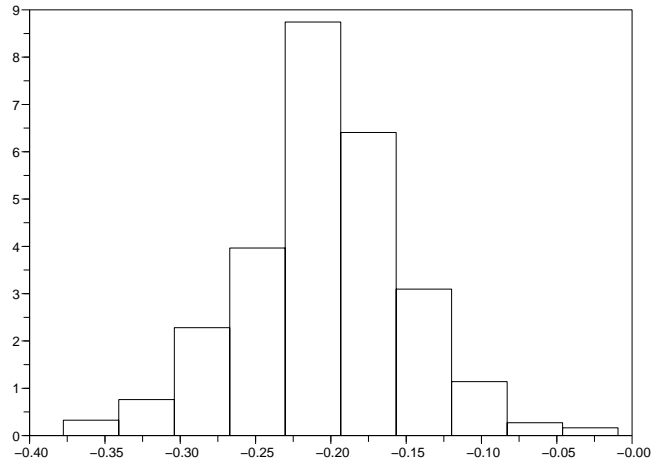


FIG. 2.10 – L’histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = -0.2$.

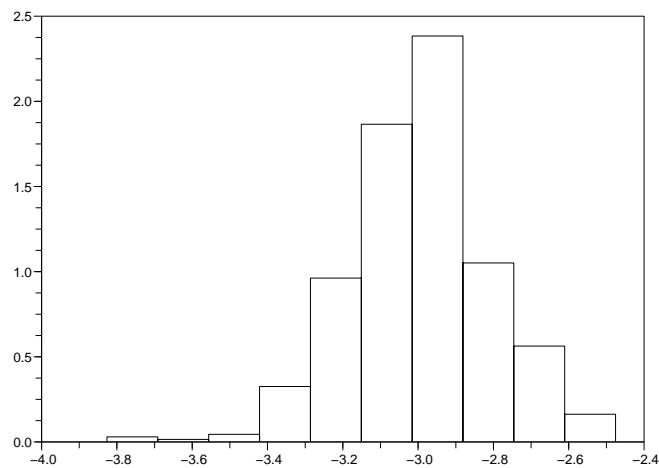


FIG. 2.11 – L’histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = -2.96$.

– Cas C : $\alpha_0 = 0.3$ et $\lambda_0 = 5.65$.

Donc, $\alpha_0 = \frac{p}{q}$ où $p = 3$ et $q = 10$ est pair. Ainsi, nous avons $\{\lambda_0\} = 0.65 \in i_0 = [\frac{6}{10}, \frac{7}{10}[$ et par suite $I_0 = [5.6, 5.7[$. La moyenne des $\hat{\alpha}_{n,i}$ est 0.3032 et leur écart-type est 0.0426. D'ailleurs, la moyenne des $\hat{\lambda}_{n,i}$ est 5.6105 et leur écart-type est 0.3459. Notons que la moyenne des estimations de λ appartient à l'intervalle I_0 .

Les figures suivantes nous donnent les histogrammes des estimations obtenues.

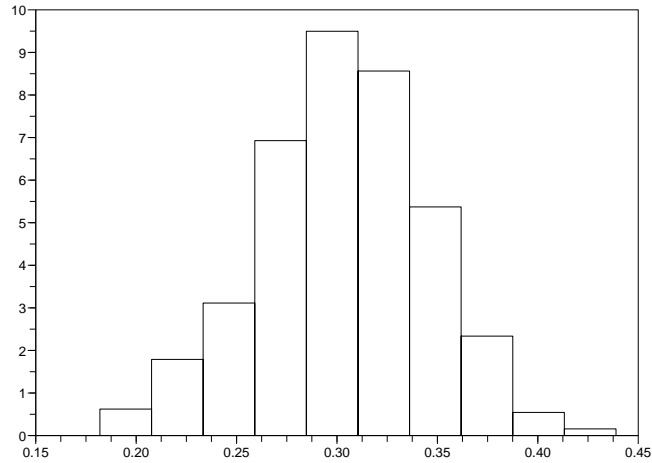


FIG. 2.12 – L'histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = 0.3$.

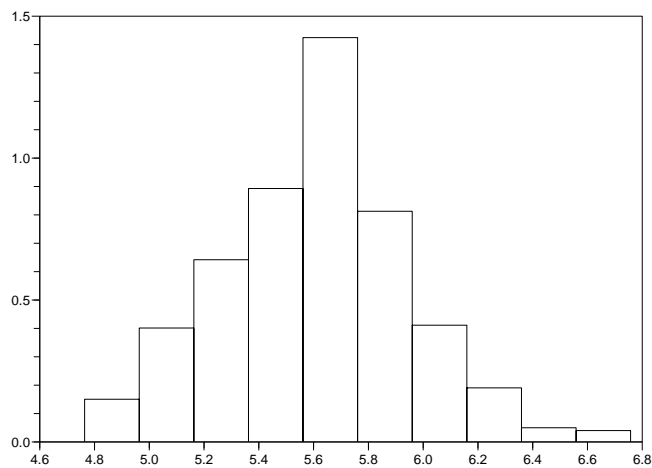


FIG. 2.13 – L'histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = 5.65$.

Rappelons que le bruit ε_t est la soustraction de deux variables aléatoires indépendantes qui suivent une loi de Poisson de paramètre μ . Ainsi, pour tout $t \in \mathbb{Z}$, nous avons

$$\mathbb{E}\varepsilon_t = \mu - \mu = 0,$$

et

$$V(\varepsilon_t) = 2\mu.$$

Dans cette section, nous avons considéré que le bruit est grand (c.à.d. $\mu \gg \frac{1}{2}$) et $|\lambda_0| \gg \frac{1}{2}$. Notons que dans ce cas, l'estimateur de Yule-Walker $\hat{\theta}_0$ des paramètres, qui constitue le point de départ de notre algorithme, est considéré comme significatif. En d'autres termes, l'estimateur des moindres carrés $\hat{\theta}_n$ calculé suivant notre algorithme doit être très proche de $\hat{\theta}_0$.

2.6 Analyse des données d'O'Donovan

Dans cette section, nous présentons une première application à des données réelles. Il s'agit de 70 observations représentant des résultats consécutifs d'un processus chimique, source : O'Donovan (1983, [55]).

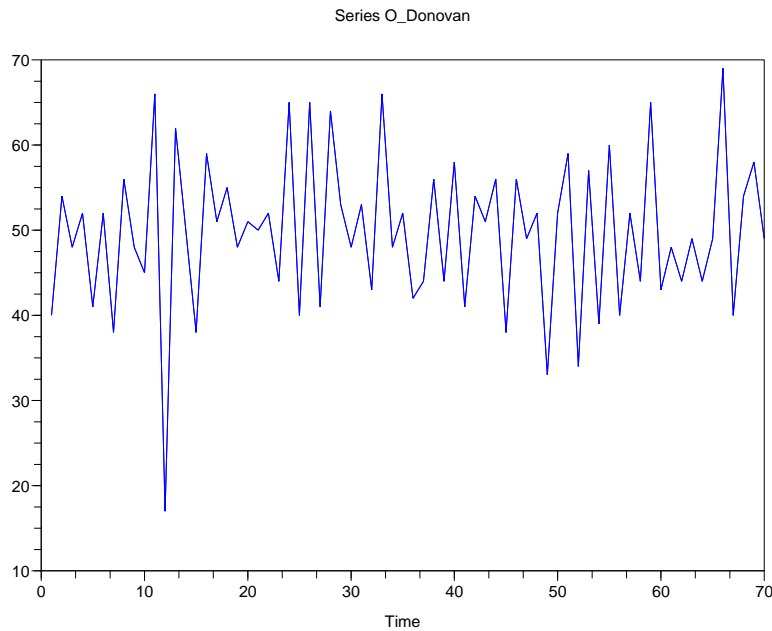


FIG. 2.14 – 70 observations provenant des résultats consécutifs d'un processus chimique, source : O'Donovan.

Notons que toutes les observations sont positives, elles varient entre 17 et 66.

De plus, la moyenne empirique est 49.6857 et la variance empirique est 84.7403.

Les Figures 2.15 et 2.16 nous donnent respectivement les fonctions d'autocorrélation empirique simples (ACF) et partielles (PACF). D'après l'ACF, nous constatons que le coefficient d'autocorrélation empirique du premier ordre est significatif. Cependant, le coefficient d'autocorrélation empirique du second ordre est presque significatif. Nous avons, $\hat{\rho}(1) = -0.588$ et $\hat{\rho}(2) = 0.272$. De plus, nous remarquons alternance des signes entre les coefficients d'autocorrélations empiriques du premier et second ordre, et les coefficients d'ordre supérieur tendent vers zéro rapidement. D'après le PACF, nous pouvons voir que seulement le coefficient d'autocorrélation partiel du premier ordre est significatif. Ainsi, O'Donovan (1983, [55]) suggère un modèle AR(1) réel. Par suite, nous proposons le processus RINAR(1) pour analyser les données. Afin de comparer les modèles proposés, nous réservons les 60 premières observations pour estimer les paramètres et les 10 dernières pour comparer les performances de prévision.

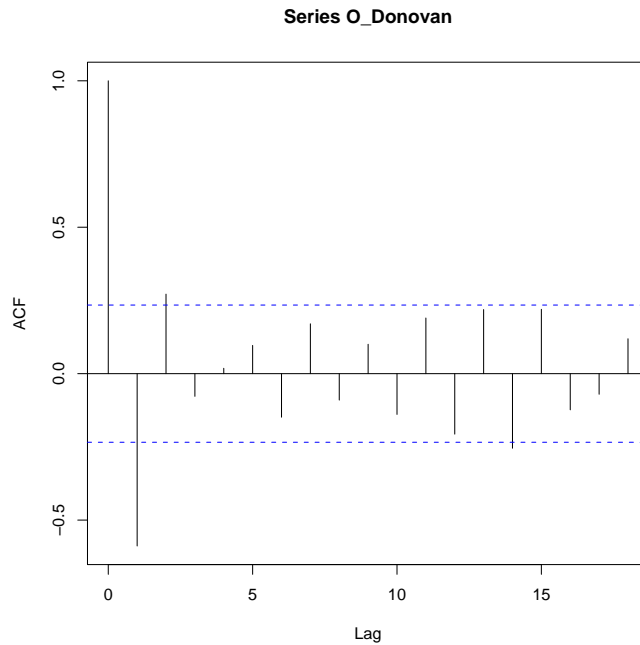


FIG. 2.15 – ACF des 70 observations d'O'Donovan.

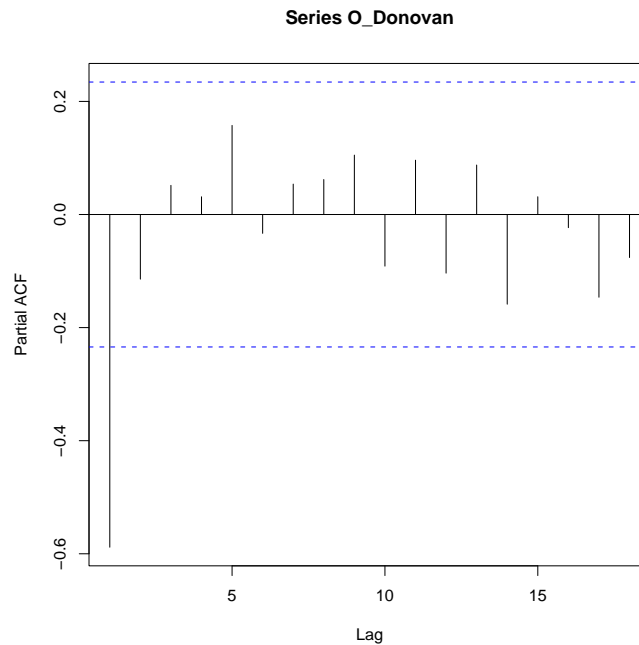


FIG. 2.16 – PACF des 70 observations d'O'Donovan.

2.6.1 Le modèle AR(1) d'O'Donovan

Pour analyser ces observations, O'Donovan (1983, [55]) propose le modèle AR(1) suivant

$$X_t = aX_{t-1} + b + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Pour estimer les paramètres du modèle, l'auteur utilise l'estimateur de Yule-Walker.

Ainsi, nous avons

$$\hat{a} = \hat{\rho}(1) = -0.6242, \quad \hat{b} = \bar{X}_n(1 - \hat{a}) = 80.5067 \quad (\text{pour les 60 premières observations}),$$

où $\hat{\rho}(1)$ est le coefficient d'autocorrélation empirique du premier ordre et \bar{X}_n est la moyenne empirique. La figure suivante montre les diagnostics basés sur les résidus obtenus tels que leur ACF et le test de Ljung-Box. Ces derniers confirment le choix du modèle AR(1).

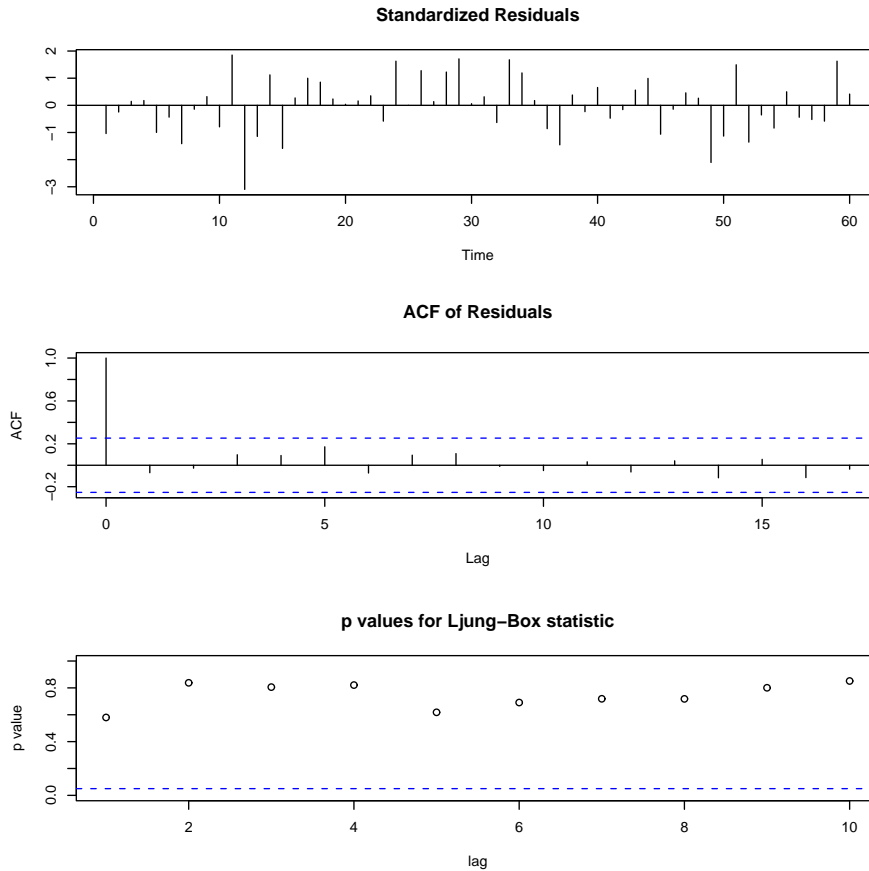


FIG. 2.17 – Test sur les résidues obtenus en utilisant AR(1) pour modéliser les données d'O'Donovan.

Soit T un instant fixé. La prévision à un pas \tilde{X}_{T+1} , basée sur la fonction de régression, est donnée par

$$\tilde{X}_{T+1} = \hat{a}X_T + \hat{b}. \quad (2.89)$$

En général, la valeur de \tilde{X}_{T+1} est réelle. Le passage au support discret de la série est facilement obtenu en arrondissant à l'entier le plus près, c.à.d.

$$\tilde{X}_{T+1} = \langle \hat{a}X_T + \hat{b} \rangle. \quad (2.90)$$

Le tableau suivant nous donne les résultats de prévision sur les 10 dernières observations de la série. Rappelons qu'à l'instant 60, nous avons $X_{60} = 43$.

Instant	61	62	63	64	65	66	67	68	69	70
Vraie valeur	48	44	49	44	49	69	40	54	58	49
Valeur de prévision	54	51	53	50	53	50	37	56	47	44

2.6.2 Ajustement à un modèle RINAR(1)

Maintenant, nous proposons le modèle RINAR(1) suivant

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Pour estimer les paramètres du modèle, nous considérons l'estimateur des moindres carrés $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n)$ défini par (2.21) et (2.22). Pour calculer $\hat{\theta}_n$, nous utilisons l'algorithme proposé dans la section 2.5.

Ainsi, nous obtenons

$$\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n) = (-0.625, 80.749).$$

Les résidus du modèle (voir Figure 2.18) sont définis par

$$\hat{\varepsilon}_t = X_t - \langle \hat{\alpha}_n X_{t-1} + \hat{\lambda}_n \rangle, \quad \forall t = 2, \dots, 60$$

La moyenne empirique des résidus est 0.0847 et leurs écart-type est 7.214.

Les figures 2.19 et 2.20 nous donnent respectivement l'ACF et le PACF des résidus obtenus. Notons que tous les coefficients ne sont pas significatifs. Par suite, les résidus peuvent être interprétés comme un bruit blanc.

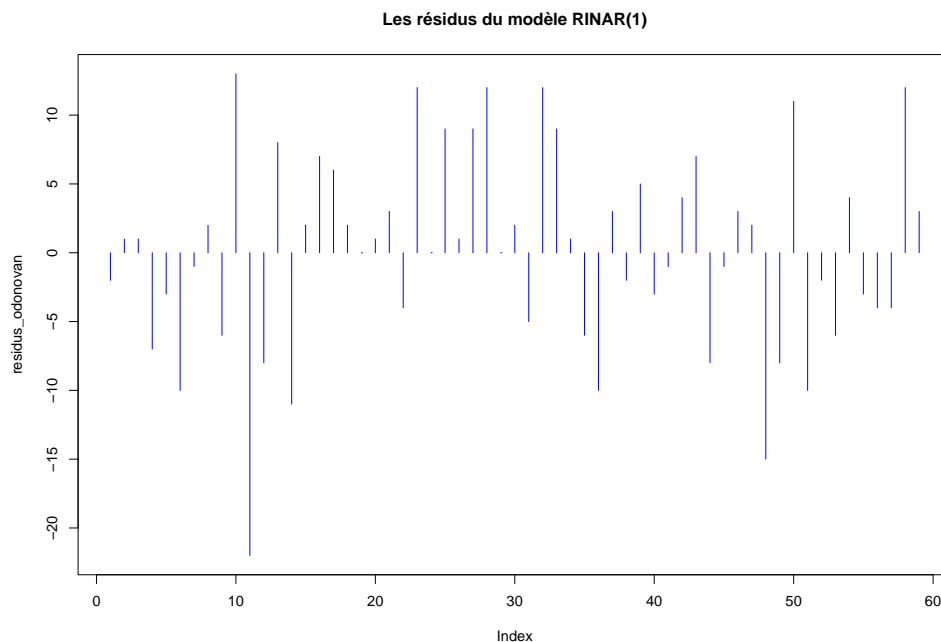


FIG. 2.18 – Les résidus obtenus en utilisant le processus RINAR(1).

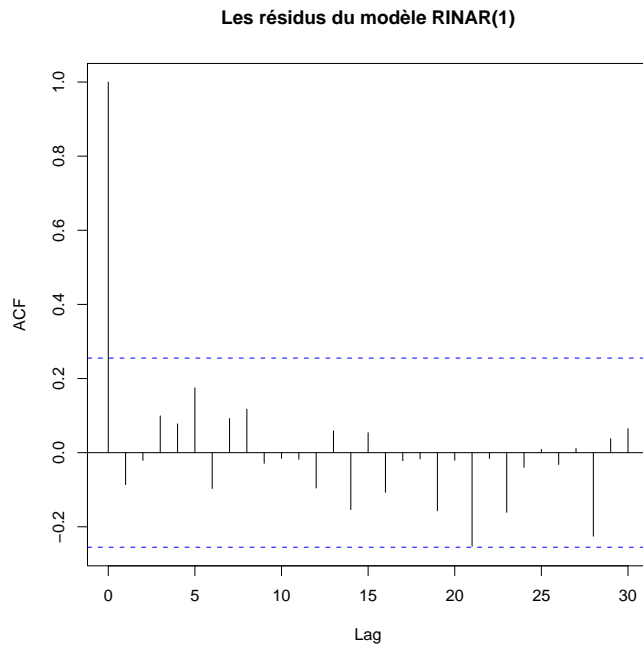


FIG. 2.19 – ACF des résidus.

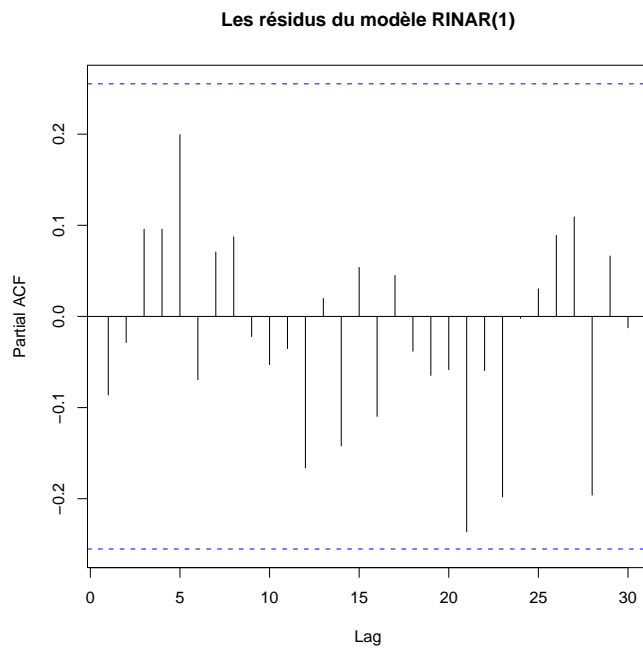


FIG. 2.20 – PACF des résidus.

La prévision à un pas \hat{X}_{T+1} , basée sur la fonction de régression, est donnée par

$$\hat{X}_{T+1} = \langle \hat{\alpha}_n X_T + \hat{\lambda}_n \rangle. \quad (2.91)$$

Notons que \hat{X}_{T+1} est une valeur entière directement. Le tableau suivant nous donne les résultats de prévision sur les 10 dernières observations de la série.

Instant	61	62	63	64	65	66	67	68	69	70
Vraie valeur	48	44	49	44	49	69	40	54	58	49
Valeur de prévision	54	51	53	50	53	50	38	56	47	44

2.6.3 Comparaison et commentaires

Une analyse basée sur des modèles stochastiques doit tenir compte de la nature entière de la série observée. Typiquement, un processus AR(1) standard ne garantit pas cette particularité.

D'autre part, comme le coefficient d'autocorrélation du premier ordre est de signe négatif, le processus INAR(1) ne peut pas modéliser la série présente.

D'ailleurs, nous notons que la valeur calculée de l'estimateur des moindres carrés $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n)$, suivant la recherche dichotomique, est très proche du point de départ de l'algorithme $\hat{\theta}_n = (\hat{a}, \hat{b})$ qui n'est autre que l'estimateur de Yule-Walker du modèle AR(1) proposé par O'Donovan.

Ceci explique les performances de prévision presque identiques des deux modèles (les prévisions sont identiques, sauf la 7^{ème} où RINAR(1) améliore la prédiction d'une unité).

2.7 Analyse des taux de variation annuels de la population Suédoise

Dans cette section, nous présentons une deuxième application à des données réelles.

Il s'agit des taux de variation (par millier d'habitants) annuels de la population Suédoise entre 1750 et 1849, notés P_t , source : Thomas (1940, [60]).

Bien que cette variation de la population doit être en général à valeurs réelles, les taux enregistrés sont des valeurs arrondies (sur une échelle de mille).

La Figure 2.21 nous montre les observations. Ainsi, nous remarquons que cette série possède des valeurs négatives. Les observations varient entre -27 et 16 , la moyenne empirique est 6.69 et la variance est 34.559 . La Figure 2.22 nous donne l'ACF et le PACF de la série. La décroissance géométrique des coefficients d'autocorrélation simple et le fait que le coefficient d'autocorrélation partiel du premier ordre est significatif justifient le choix d'un processus AR(1) pour modéliser cette série.

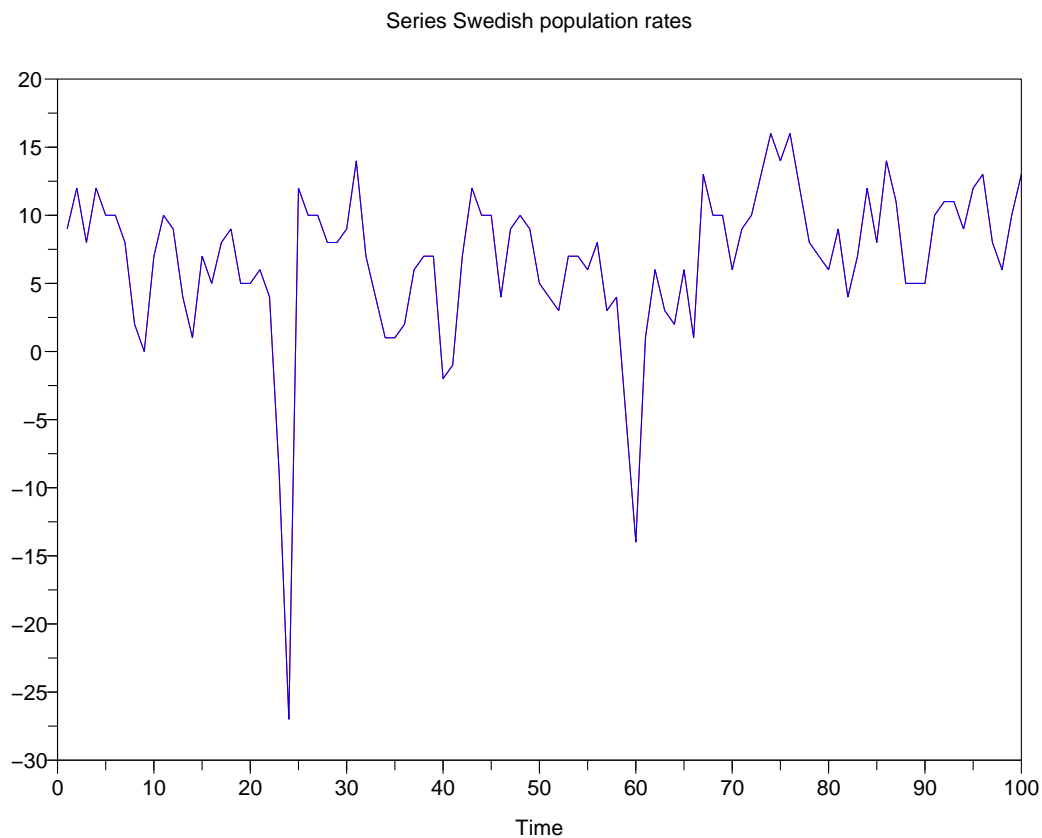


FIG. 2.21 – Les taux de variation annuels de la population Suédoise, 1750 – 1849.

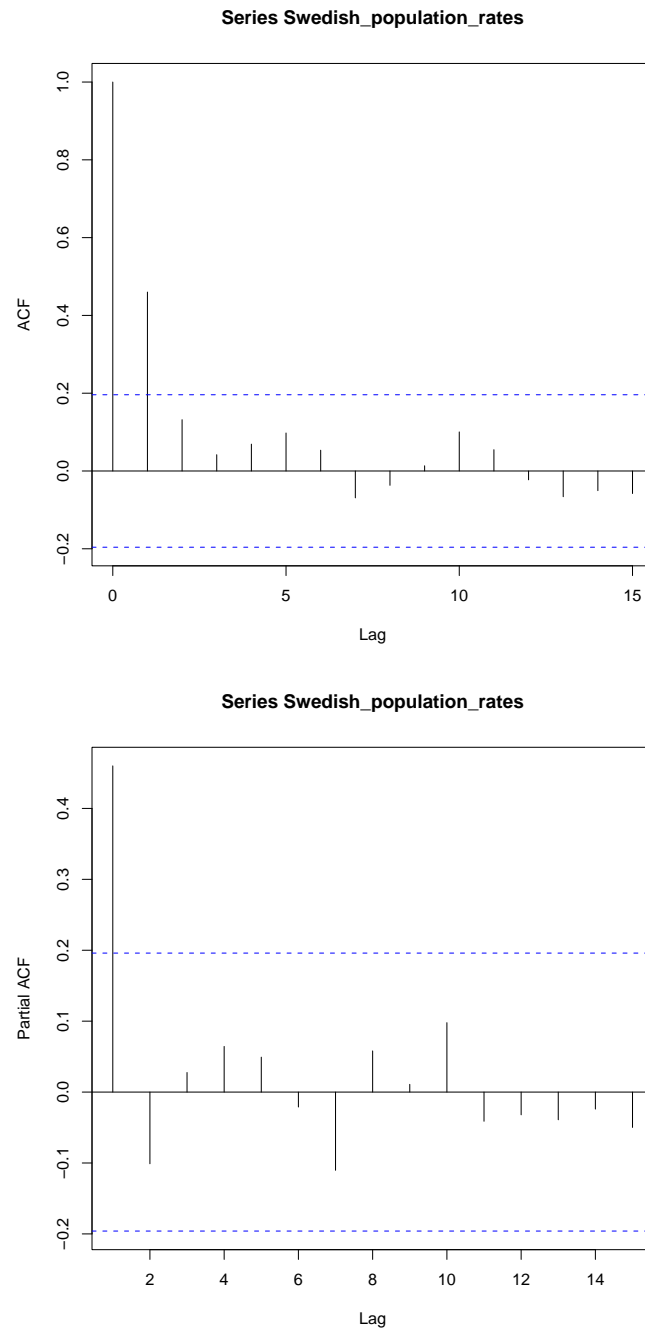


FIG. 2.22 – ACF et PACF des taux de variation annuels de la population Suédoise.

Afin de comparer les modèles proposés pour analyser cette série, nous réservons les 80 premières observations pour estimer les paramètres et les 20 dernières pour comparer les performances de prévision.

2.7.1 Ajustement à un modèle AR(1)

En premier lieu, nous proposons le modèle AR(1) suivant

$$P_t = a P_{t-1} + b + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Pour estimer les paramètres du modèle, nous utilisons l'estimateur de Yule-Walker. Ainsi,

$$\hat{a} = \hat{\rho}(1) = 0.4505, \quad \hat{b} = \bar{P}_n(1 - \hat{a}) = 3.337 \quad (\text{pour les 80 premières observations}),$$

où $\hat{\rho}(1)$ est le coefficient d'autocorrélation empirique du premier ordre et \bar{P}_n est la moyenne empirique.

La prévision à un pas, à l'instant $T + 1$, basée sur la fonction de régression et suivie d'une étape finale d'arrondi, est égale

$$\tilde{P}_{T+1} = \langle \hat{a}P_T + \hat{b} \rangle.$$

Pour les 20 dernières observations de la série, l'erreur quadratique (resp. absolue) moyenne des prévisions est donnée par

$$EQM = \frac{1}{20} \sum_{t=81}^{100} (P_t - \tilde{P}_t)^2 = 5.8, \quad EAM = \frac{1}{20} \sum_{t=81}^{100} |P_t - \tilde{P}_t| = 2.1.$$

2.7.2 Ajustement à un modèle RINAR(1)

Maintenant, dans le but d'analyser cette série, nous proposons le modèle RINAR(1) suivant

$$P_t = \langle \alpha P_{t-1} + \lambda \rangle + \varepsilon_t, \quad \forall t \in \mathbb{Z}.$$

Pour estimer les paramètres du modèle, nous considérons l'estimateur des moindres carrés $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n)$ défini par (2.21) et (2.22). Pour calculer $\hat{\theta}_n$, nous utilisons l'algorithme proposé dans la section 2.5. Ainsi, nous obtenons

$$\hat{\theta}_n = (\hat{\alpha}_n, \hat{\lambda}_n) = (0.469, 3.559).$$

Les résidus du modèle (voir Figure 2.23) sont définis par

$$\hat{\varepsilon}_t = P_t - \langle \hat{\alpha}_n P_{t-1} + \hat{\lambda}_n \rangle, \quad \forall t = 2, \dots, 80$$

La moyenne empirique des résidus est -0.265 et leur écart-type est 5.553 .

La figure (2.24) nous donne l'ACF et le PACF des résidus obtenus.

Notons que tous les coefficients ne sont pas significatifs. Par suite, les résidus peuvent être interprétés comme un bruit blanc.

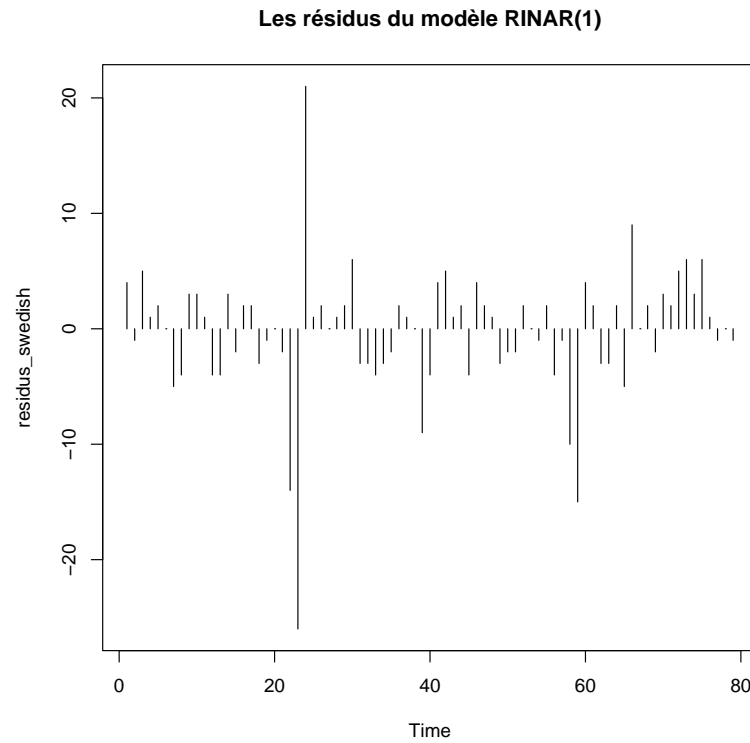


FIG. 2.23 – Les résidus obtenus en utilisant le processus RINAR(1).

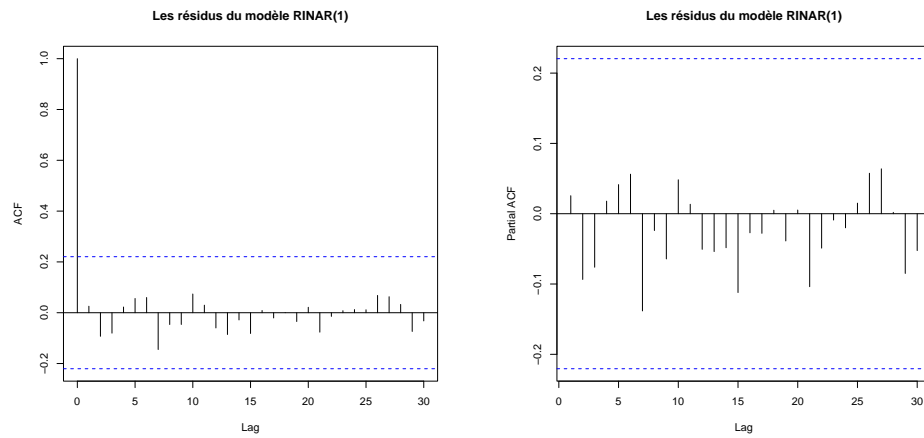


FIG. 2.24 – ACF et PACF des résidus.

La prévision à un pas, à l'instant $T + 1$, basée sur la fonction de régression, est égale

$$\hat{P}_{T+1} = \langle \hat{\alpha}_n P_T + \hat{\lambda}_n \rangle.$$

Pour les 20 dernières observations de la série, l'erreur quadratique (resp. absolue) moyenne des prévisions est donnée par

$$EQM = \frac{1}{20} \sum_{t=81}^{100} \left(P_t - \hat{P}_t \right)^2 = 4, \quad EAM = \frac{1}{20} \sum_{t=81}^{100} |P_t - \hat{P}_t| = 1.7.$$

2.7.3 L'approche de McCleary & Hay

McCleary & Hay (1980, [49]) ont analysé à la fois la série d'indice de récolte de la population suédoise (Swedish Harvest Index) et les taux de variation de la population entre 1750 et 1849. En fait, l'indice de récolte est considéré comme une mesure brute de la production alimentaire, a une forte influence sur les taux de variation de la population.

Pour ces deux séries, les auteurs ont proposé un modèle MA(1) réel. Ainsi, pour la série présente le modèle proposé est

$$P_t - \mu = \beta_1 \varepsilon_{t-1} + \varepsilon_t,$$

avec μ et β_1 deux paramètres réels. Pour estimer ces derniers nous utilisons le logiciel R. Ainsi, nous obtenons

$$\hat{\beta}_1 = 0.4307 \text{ et } \hat{\mu} = 6.075 \text{ (la moyenne empirique des 80 premières observations)}$$

Soit $\mathcal{F}_T = \sigma \{P_T, P_{T-1}, \dots\}$. La prévision à un pas, à l'instant $T + 1$, basée sur la fonction de régression (espérance conditionnelle) et suivie d'une étape finale d'arrondi, est égale à

$$\check{P}_{T+1} = \mathbb{E} [P_{T+1} \mid \mathcal{F}_T] = \langle \hat{\mu} \rangle = 6.$$

Pour les 20 dernières observations de la série, l'erreur quadratique (resp. absolue) moyenne des prévisions est donnée par

$$EQM = \frac{1}{20} \sum_{t=81}^{100} \left(P_t - \check{P}_t \right)^2 = \frac{1}{20} \sum_{t=81}^{100} (P_t - 6)^2 = 18.75,$$

$$EAM = \frac{1}{20} \sum_{t=81}^{100} |P_t - \check{P}_t| = \frac{1}{20} \sum_{t=81}^{100} |P_t - 6| = 3.65.$$

2.7.4 Comparaison et commentaires

Toute analyse de cette série, basée sur des modèles stochastiques doit tenir compte de la nature entière des observations. Typiquement, un processus AR(1) (resp. MA(1)) standard ne garantit pas cette particularité.

D'autre part, comme cette série possède des observations négatives, aucun processus basé sur l'opérateur d'aminicissement peut modéliser la série (en particulier INAR(1) et INMA(1)).

La Figure 2.25 montre les résultats de prévision de RINAR(1) et AR(1) pour les 20 dernières observations de la série.

Comparées aux performances du modèle AR(1), les prévisions du processus RINAR(1) sont différentes en 8 points et respectent plus les variations de la série. De plus, notons que l'erreur quadratique (resp. absolue) moyenne décroît de 5.8 (resp. 2.1) à 4 (resp. 1.7).

D'ailleurs, les résultats de prévision du modèle RINAR(1) sont bien meilleurs que ceux du modèle MA(1) proposé par McCleary & Hay, l'erreur quadratique (resp. absolue) moyenne sur les 20 dernières observations décroît de 18.75 (resp. 3.6) à 4 (resp. 1.7).

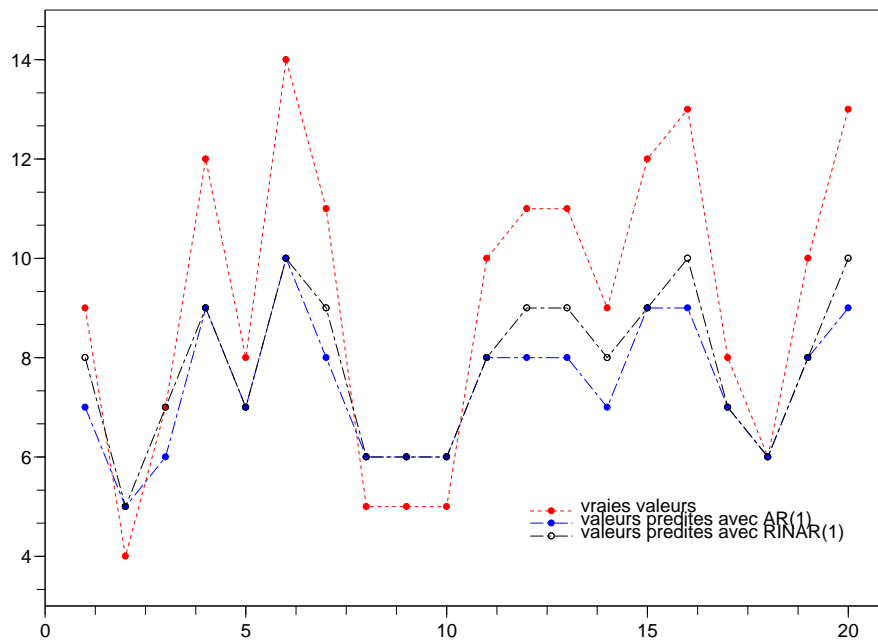


FIG. 2.25 – Les 20 dernières observations des taux de variation de la population Suédoise et leurs prévisions basées sur les modèles RINAR(1) et AR(1).

Chapitre 3

Le modèle RINAR(p)

Sommaire

3.1	Stationnarité et ergodicité du modèle RINAR(p)	80
3.2	Estimation des paramètres	84
3.2.1	La consistance forte de l'estimateur des moindres carrés quand un au moins des coefficients de régression α_j^* est irrationnel	91
3.2.2	La consistance forte de l'estimateur des moindres carrés quand tous coefficients de régression α_j^* sont rationnels	95
3.3	Méthode numérique pour calculer $\hat{\theta}_n$	104
3.3.1	Initialisation	104
3.3.2	Recherche dichotomique successive	104
3.3.3	Une expérience de simulation	106
3.4	Analyse des données de Fürth	110
3.4.1	Le modèle INAR(2) de Jung & Termayne	111
3.4.2	Ajustement à un modèle RINAR(2)	112

Dans ce chapitre nous examinons le modèle arrondi à valeurs entières autorégressif d'ordre p , noté RINAR(p). Ce dernier est une extension directe et naturelle du processus RINAR(1) examiné dans le Chapitre 2. Notons que, le modèle RINAR(p) a été étudié par Kachour (2009, [42]).

Dans le premier paragraphe, nous donnons les conditions de stationnarité et d'ergodicité du processus. Dans le deuxième paragraphe, nous proposons l'estimateur des moindres carrés pour estimer les paramètres du modèle RINAR(p). Puis, nous examinons le problème d'identifiabilité, dû à l'opérateur d'arrondi, et la consistance de ce dernier.

Dans le troisième paragraphe, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Finalement, dans le quatrième paragraphe, nous présentons une application réelle traitée avec un modèle RINAR(p). Notons que dans ce chapitre nous gardons les mêmes notations et définitions que dans le Chapitre 2.

3.1 Stationnarité et ergodicité du modèle RINAR(p)

Définition 3.1.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Le processus $\{X_t, t \in \mathbb{Z}\}$ est dit un processus RINAR(1) si pour tout t , il possède la présentation suivante

$$X_t = \langle \sum_{j=1}^p \alpha_j X_{t-j} + \lambda \rangle + \varepsilon_t, \quad (3.1)$$

où $\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près, (ε_t) est une suite de variables aléatoires i.i.d. centrées à valeurs dans \mathbb{Z} , définie sur $(\Omega, \mathcal{A}, \mathbb{P})$, λ et les α_j sont des paramètres réels.

Ce modèle représente une extension naturelle du processus RINAR(1) étudié dans le Chapitre 2. Ainsi, notons que l'opérateur d'arrondi peut être interprété comme une fonction de censure sur le modèle AR(p) et λ comme la moyenne du bruit non-centré ($\varepsilon'_t = \varepsilon_t + \lambda$).

L'étude du processus RINAR(p) peut être effectuée par le biais du processus vectoriel suivant

$$Y_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix} = \begin{pmatrix} \langle \sum_{j=1}^p \alpha_j X_{t-j} + \lambda \rangle \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (3.2)$$

Comme le bruit (ε_t) est suite de variables aléatoires i.i.d., le processus (Y_t) défini par (3.2), forme une chaîne de Markov homogène avec un espace d'états $E = \mathbb{Z}^p$ et une probabilité de transition

$$\pi(x, y) = \mathbb{P}(\varepsilon_1 = y_1 - \langle \sum_{j=1}^p \alpha_j x_j + \lambda \rangle) \mathbb{1}_{y_2=x_1, \dots, y_p=x_{p-1}}, \quad \forall x = (x_j), y = (y_j) \in E. \quad (3.3)$$

Rappelons que, pour tout $x = (x_1, \dots, x_p) \in \mathbb{R}^p$, nous avons $\|x\|_1 = |x_1| + \dots + |x_p|$.

De plus, pour toute mesure μ et toute fonction g sur E , nous posons

$$\mu(g) = \int g(x) d\mu(x) = \sum_{i \in E} g(i) \mu(i).$$

D'autre part, pour tout $n \in \mathbb{N}^*$, nous avons

$$\begin{aligned} (\pi^n g)(x) &= \mathbb{E}[g(X_{t+n}) \mid X_t = x], \quad \forall t \in \mathbb{Z} \\ &= \mathbb{E}_x[g(X_n)], \end{aligned}$$

où $\mathbb{E}_x[\cdot]$ représente l'espérance conditionnelle $\mathbb{E}[\cdot \mid X_0 = x]$.

Proposition 3.1.1. *Supposons que :*

1. *La chaîne de Markov (Y_t) est irréductible ;*
2. *pour un certain $k > 1$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;*
3. $\sum_{j=1}^p |\alpha_j| < 1$.

Alors,

1. *(Y_t) possède une unique mesure de probabilité invariante, notée μ ayant un moment d'ordre k .*
2. *Pour tout $x \in E$ et $f \in L^1(\mu)$ nous avons*

$$\frac{1}{n} \sum_{k=1}^n f(Y_k) \longrightarrow \mu(f), \quad \mathbb{P}_x \text{ p.s.}$$

où \mathbb{P}_x représente la probabilité conditionnelle $\mathbb{P}(\cdot \mid Y_0 = x)$.

Preuve. Tout d'abord, nous définissons sur E les fonctions φ et V par

$$x \mapsto \varphi(x) = \begin{pmatrix} |x_1| \\ \vdots \\ |x_p| \end{pmatrix}, \quad x \mapsto V(x) = \left(\sum_{j=1}^p |x_j| \right)^k = (\|x\|_1)^k.$$

Comme V est positive et $\lim_{|x| \rightarrow \infty} V(x) = \infty$, V est donc une fonction de Lyapunov.

Soit " \leq " une relation d'ordre partiel sur \mathbb{R}^p définie par

$$x \leq y \iff |x_j| \leq |y_j|, \quad \forall j = 1, \dots, p,$$

pour $x = (x_j)$ et $y = (y_j)$. Soient

$$A = \begin{pmatrix} |\alpha_1| \cdots |\alpha_p| \\ I_{p-1} & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} |b| \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

où I_{p-1} est la matrice d'identité d'ordre $p-1$ et b est une constante.

Ainsi, nous avons

$$x \leq y \implies \begin{cases} A \varphi(x) + B \leq A \varphi(y) + B, \\ \|x\|_1 = \|\varphi(x)\|_1 \leq \|\varphi(y)\|_1 = \|y\|_1. \end{cases}$$

Rappelons que, pour tout $a \in \mathbb{R}$, nous avons

$$|\langle a \rangle| = \langle |a| \rangle \leq |a| + \frac{1}{2}.$$

Par conséquent, nous trouvons

$$|X_t| = |\langle \sum_{j=1}^p \alpha_j X_{t-j} + \lambda \rangle + \varepsilon_t| \leq \sum_{j=1}^p |\alpha_j| |X_{t-j}| + |\lambda| + |\varepsilon_t| + \frac{1}{2}.$$

Soit

$$\zeta_t = \begin{pmatrix} |\lambda| + |\varepsilon_t| + \frac{1}{2} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

Donc, nous obtenons

$$\varphi(Y_t) \leq A \varphi(Y_{t-1}) + \zeta_t,$$

Ainsi pour tout $n \in \mathbb{N}^*$, nous trouvons par itération

$$\varphi(Y_{t+n}) \leq A^n \varphi(Y_t) + A^{n-1} \zeta_{t+1} + A^{n-2} \zeta_{t+2} + \cdots + A \zeta_{t+n-1} + \zeta_{t+n}.$$

Par suite,

$$\begin{aligned} \|\varphi(Y_{t+n})\|_1 &\leq \|A^n \varphi(Y_t) + A^{n-1} \zeta_{t+1} + \cdots + A \zeta_{t+n-1} + \zeta_{t+n}\|_1 \\ &\leq \|A^n \varphi(Y_t)\|_1 + \|A^{n-1} \zeta_{t+1}\|_1 + \cdots + \|A \zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1 \\ &\leq \|A^n \varphi(Y_t)\|_1 + \|A^{n-1}\|_1 \|\zeta_{t+1}\|_1 + \cdots + \|A\|_1 \|\zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1, \end{aligned}$$

où $\|\cdot\|_1$ est la norme matricielle associée à la norme $\|\cdot\|_1$. Il en découle,

$$\begin{aligned} V(Y_{t+n}) &= (\|Y_{t+n}\|_1)^k = (\|\varphi(Y_{t+n})\|_1)^k \\ &\leq (\|A^n \varphi(Y_t)\|_1 + \|A^{n-1}\|_1 \|\zeta_{t+1}\|_1 + \cdots + \|A\|_1 \|\zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1)^k. \end{aligned}$$

Donc,

$$((\pi^n V)(x))^{\frac{1}{k}} \leq \left(\mathbb{E}_x [\|A^n \varphi(Y_0)\|_1 + \|A^{n-1}\|_1 \|\zeta_{t+1}\|_1 + \cdots + \|A\|_1 \|\zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1]^k \right)^{\frac{1}{k}}$$

Notons que, pour tout $j \in \{1, \dots, n\}$, $\|\zeta_{t+j}\|_1 = |\varepsilon_{t+j}| + |\lambda| + \frac{1}{2}$.

Ainsi, comme $\mathbb{E}|\varepsilon_t|^k < \infty$, nous avons

$$b = \left(\mathbb{E} [\|\zeta_{t+j}\|_1^k] \right)^{\frac{1}{k}} < \infty, \quad \forall j \in \{1, \dots, n\}.$$

Par conséquent, d'après l'inégalité de Minkowski, nous obtenons

$$\begin{aligned} ((\pi^n V)(x))^{\frac{1}{k}} &\leq \left(\mathbb{E}_x \|A^n \varphi(Y_0)\|_1^k \right)^{\frac{1}{k}} + S_n(A) b \\ &= \|A^n \varphi(x)\|_1 + S_n(A) b \\ &\leq \|A^n\|_1 \|x\|_1 + S_n(A) b, \end{aligned}$$

où

$$S_n(A) = \|A^{n-1}\|_1 + \dots + \|A\|_1 + 1,$$

Notons que, pour toute norme matricielle $\|\cdot\|$, nous avons $\|A^n\| \xrightarrow{\frac{1}{n}} \rho(A)$ quand $n \rightarrow \infty$, où $\rho(A)$ est le rayon spectral de A .

Sous l'hypothèse $\sum_{j=1}^p |\alpha_j| < 1$, nous avons $\rho(A) < 1$. Donc, il existe $n_0 \in \mathbb{N}^*$ tel que $\forall N \geq n_0$, nous avons $\|A^N\| = \alpha' < 1$. Par suite, d'après le critère de Cauchy, nous obtenons que la série $\sum \|A^i\|_1$ est convergente.

Ainsi, nous trouvons

$$((\pi^N V)(x))^{\frac{1}{k}} \leq \alpha' \|x\|_1 + \beta = \|x\|_1 \left(\alpha' + \frac{\beta}{\|x\|_1} \right)$$

et par suite

$$\frac{(\pi^N V)(x)}{V(x)} \leq \left(\alpha' + \frac{\beta}{\|x\|_1} \right)^k,$$

où $\alpha' < 1$ et $0 < \beta < \infty$. Par conséquent, il existe $N \in \mathbb{N}^*$ ($N \geq n_0$) tel que

$$\limsup_{\|x\|_1 \rightarrow \infty} \frac{\pi^N V(x)}{V(x)} \leq |\alpha'|^k < 1.$$

Pour tout $x_0 \in E = \mathbb{Z}^p$ nous définissons la mesure empirique par :

$$\mu_n(\cdot) = \frac{1}{n} [\pi^1(x_0, \cdot) + \dots + \pi^n(x_0, \cdot)],$$

où π^n représente la probabilité de transition de la chaîne de Markov (Y_t) en n pas.

Finalement, nous obtenons

$$\limsup_{n \rightarrow \infty} \mu_n V \leq l < \infty.$$

La chaîne de Markov (Y_t) vérifie le critère de Lyapunov, avec V comme fonction de Lyapunov. Ainsi, (μ_n) est une suite tendue et pour tout point limite μ , nous avons $\mu V \leq l < \infty$

(voir Duflo [22], proposition 2.1.6). Alors, μ est une mesure π -invariante. De plus, nous avons que la chaîne de Markov (Y_t) est irréductible. Il en résulte que cette dernière soit récurrente positive et par conséquent μ est l'unique mesure invariante, avec $\mu V < \infty$. La conclusion 2 de la proposition est une conséquence directe du théorème ergodique classique pour les chaînes de Markov. \square

Notons que l'hypothèse 3 est équivalente à la condition nécessaire qui assure la stationnarité du modèle AR(p) réel. D'autre part, notons que le support de la mesure de probabilité invariante est l'espace des états $E = \mathbb{Z}^p$ de la chaîne (Y_t) . En d'autres termes, pour tout $x \in \mathbb{Z}^p$, nous avons $\mu(x) > 0$.

3.2 Estimation des paramètres

Tout d'abord, nous donnons quelques notations rappels et définitions.

Soit $x = (x_1, \dots, x_p)^\tau \in \mathbb{R}^p$. Rappelons que $\|x\|_1 = \sum_{j=1}^p |x_j|$. Nous associons à la norme $\|\cdot\|_1$ sur \mathbb{R}^p , une norme sur $(\mathbb{R}^p)^2$ définie par

$$|z| = \|y_1\|_1 + \|y_2\|_1, \quad \forall z = (y_1, y_2) \in (\mathbb{R}^p)^2.$$

Nous définissons l'arrondi du vecteur x , noté $\langle x \rangle$, par

$$\langle x \rangle = (\langle x_1 \rangle, \dots, \langle x_p \rangle)^\tau \in E = \mathbb{Z}^p.$$

Soit $\theta = (\alpha_1, \dots, \alpha_p, \lambda) \in \Theta \subset]-1, 1]^p \times \mathbb{R}$. Par la suite, nous considérons que l'espace des paramètres Θ est compact. Nous définissons la fonction de régression, notée f , par

$$f(x; \theta) = f(x_1, \dots, x_p; \theta) = \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle. \quad (3.4)$$

Ainsi, le modèle RINAR(p) peut être écrit sous la forme suivante

$$Y_t = \begin{pmatrix} f(Y_{t-1}; \theta) \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (3.5)$$

Soient

$$M = \begin{pmatrix} \alpha_1 & \dots & \alpha_p \\ I_{p-1} & 0 \end{pmatrix}, \quad \xi = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{et} \quad \eta_t = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (3.6)$$

où I_{p-1} est la matrice identité d'ordre $p-1$. Finalement, $\text{RINAR}(p)$ s'écrit de la manière suivante

$$Y_t = F(Y_{t-1}; \theta) + \eta_t, \quad (3.7)$$

avec

$$F(x; \theta) = \langle M \ x + \xi \rangle, \ \forall \ x \in E = \mathbb{Z}^p. \quad (3.8)$$

Soient $\theta_0 = (\alpha_1^*, \dots, \alpha_p^*, \lambda^*)$ la vraie valeur du paramètre θ et \mathbb{P}_{θ_0} la distribution de la probabilité de la chaîne (Y_t) sous le vrai modèle. De plus, toute convergence $\xrightarrow{p.s.}$ signifie une convergence $p.s.$ sous $\mathbb{P}_{\theta_0, x}$, ce qui se tient indépendamment de l'état initial x .

Soient $X_{-p+1}, \dots, X_0, \dots, X_n$ des observations du processus $\text{RINAR}(p)$.

Pour l'estimation du paramètre θ , nous considérons l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \varphi_n(\theta), \quad (3.9)$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n (X_t - f(Y_{t-1}; \theta))^2 = \frac{1}{n} \sum_{t=1}^n (\|Y_t - F(Y_{t-1}; \theta)\|_1)^2. \quad (3.10)$$

Soit l'hypothèse suivante.

Hypothèse [H2]

1. Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible ;
2. Pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;
3. $\sum_{j=1}^p |\alpha_j^*| < 1$, où α_j^* est la vraie valeur de α_j pour tout $j \in \{1, \dots, p\}$;
4. L'espace des paramètres Θ , est un compact sous-ensemble de $(]-1, 1])^p \times \mathbb{R}$.

Supposons que [H2] soit vérifiée. Donc, sous \mathbb{P}_{θ_0} et d'après la Proposition 3.1.1, (Y_t) possède une unique mesure invariante μ_{θ_0} telle que $\mu_{\theta_0}(| \cdot |^k) < \infty$ avec $k \geq 2$.

Il en résulte que la chaîne double (Z_t) avec $Z_t = (Y_{t-1}, Y_t)$ possède des propriétés similaires. En particulier, la chaîne (Z_t) possède aussi une unique mesure invariante $\sigma_{\theta_0} = \mu_{\theta_0} \otimes \pi_{\theta_0}$, c.à.d.

$$\sigma_{\theta_0}(z) = \mu_{\theta_0}(x) \pi_{\theta_0}(x, y), \ \forall \ z = (x, y) \in E^2 = (\mathbb{Z}^p)^2,$$

où π_{θ_0} est la probabilité de transition de la chaîne (Y_t) sous le vrai modèle.

De plus, comme $\mu_{\theta_0}(\| \cdot \|_1^k) < \infty$, il s'ensuit $\sigma_{\theta_0}(| \cdot |^k) < \infty$, où $| \cdot |$ est la norme sur $(\mathbb{R}^p)^2$

associée à la norme $\|\cdot\|_1$ (définie sur \mathbb{R}^p).

Soient les fonctions :

$$g_\theta(z) = (\|y - F(x; \theta)\|_1)^2, \quad z = (x, y) \in E^2, \quad \theta \in \Theta, \quad (3.11)$$

et

$$K(\theta) = \sigma_{\theta_0} g_\theta, \quad \theta \in \Theta. \quad (3.12)$$

Soit \mathbb{P}_n la mesure empirique générée par les observations Z_1, \dots, Z_n

$$\mathbb{P}_n(z) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Z_i=z}, \quad z = (x, y) \in E^2.$$

Par conséquent, d'après (3.10), la fonction de contraste φ_n est égale

$$\varphi_n(\theta) = \mathbb{P}_n g_\theta \quad (3.13)$$

et par suite,

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \mathbb{P}_n g_\theta. \quad (3.14)$$

Lemme 3.2.1. *Pour tout $\theta \in \Theta$, nous avons*

$$g_\theta \in L^1(\sigma_{\theta_0}).$$

Preuve. Rappelons que pour tout $a \in \mathbb{R}$, nous avons $|\langle a \rangle| = \langle |a| \rangle \leq |a| + \frac{1}{2}$.

Dans la preuve suivante, nous notons par c une constante générique dont la valeur exacte peut changer pendant le développement mathématique.

Pour tout $x = (x_1, \dots, x_p)^\tau \in E$, nous avons

$$F(x; \theta) = \langle M x + \xi \rangle = \begin{pmatrix} f(x; \theta) \\ \langle x_1 \rangle \\ \vdots \\ \langle x_{p-1} \rangle \end{pmatrix},$$

où

$$f(x; \theta) = \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle.$$

Comme Θ est compact, il existe $B > 0$ tel que $|\lambda| \leq B$. Ainsi, nous obtenons

$$\begin{aligned}
\|F(x; \theta)\|_1 &= |f(x; \theta)| + |\langle x_1 \rangle| + \cdots + |\langle x_{p-1} \rangle| \\
&= \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle + \langle |x_1| \rangle + \cdots + \langle |x_p| \rangle \\
&\leq (1 + |\alpha_1|) |x_1| + \cdots + (1 + |\alpha_{p-1}|) |x_{p-1}| + |\alpha_p| |x_p| + |\lambda| + \frac{p}{2} \\
&\leq (1 + |\alpha_1|) |x_1| + \cdots + (1 + |\alpha_{p-1}|) |x_{p-1}| + |\alpha_p| |x_p| + B + \frac{p}{2} \\
&\leq c (1 + \|x\|_1).
\end{aligned}$$

Par suite, pour tout $z = (x, y) \in E^2$, nous avons

$$\begin{aligned}
\|y - F(x; \theta)\|_1 &\leq \|y\|_1 + \|F(x; \theta)\|_1 \\
&\leq c (1 + \|x\|_1 + \|y\|_1) \\
&= c (1 + |z|),
\end{aligned}$$

où $|\cdot|$ est la norme sur $(\mathbb{R}^p)^2$ associée à la norme $\|\cdot\|_1$ définie sur \mathbb{R}^p .

Il en résulte, pour tout $z \in E^2$,

$$g_\theta(z) = (\|y - F(x; \theta)\|_1)^2 \leq c (1 + |z|^2). \quad (3.15)$$

Comme $\sigma_{\theta_0}(|\cdot|^k) < \infty$, où $k \geq 2$, nous trouvons

$$g_\theta \in L^1(\sigma_{\theta_0}).$$

□

La proposition suivante nous donne la limite de la fonction de contraste φ_n , où la convergence se tient uniformément sur Θ .

Proposition 3.2.1. *Supposons que [H2] soit vérifiée. Donc, pour tout $\theta \in \Theta$,*

1. $\varphi_n(\theta) \xrightarrow{p.s.} K(\theta)$;
2. $K(\theta) - K(\theta_0) = \mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right)$.

De plus, nous avons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0.$$

Preuve. — La première conclusion est simple à vérifier. Comme $\mathbb{P}_n \xrightarrow{p.s.} \sigma_{\theta_0}$ et d'après le Lemme 3.2.1, $g_\theta \in L^1(\sigma_{\theta_0})$, nous obtenons

$$\varphi_n(\theta) = \mathbb{P}_n g_\theta \xrightarrow{p.s.} \sigma_{\theta_0} g_\theta = K(\theta).$$

– Maintenant, notre objectif est de montrer que la fonction $K(\theta)$ vérifie

$$K(\theta) - K(\theta_0) = \mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right).$$

Pour cela, nous allons montrer que

$$\varphi_n(\theta) - \varphi_n(\theta_0) \xrightarrow{p.s.} \mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right).$$

Les notations et les définitions suivantes seront utilisées dans le reste de cette preuve.

Soit $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$ la filtration naturelle associée au processus RINAR(p),

où $\mathcal{F}_n = \sigma(\varepsilon_t, 0 \leq t \leq n)$ pour tout $n \in \mathbb{N}^*$, et \mathcal{F}_0 est la σ -algèbre dégénérée.

Si (M_n) est une martingale de carré intégrable par rapport à \mathcal{F} , nous notons par $([M]_n)$ son processus croissant défini par

$$[M]_0 = 0, \quad [M]_n = [M]_{n-1} + \mathbb{E}(\| [M]_n - [M]_{n-1} \|_1^2 \mid \mathcal{F}_{n-1}) \quad \text{pour } n \geq 1.$$

Soit $[M]_\infty = \lim [M]_n$. Sur $\{[M]_\infty < \infty\}$, nous avons $M_n \xrightarrow{p.s.} M_\infty$, où M_∞ est une variable aléatoire finie. En revanche, sur $\{[M]_\infty = \infty\}$, nous obtenons $\frac{M_n}{[M]_n} \xrightarrow{p.s.} 0$ (voir Duflo [22], Théorème 1.3.15, p. 20).

Rappelons que, sous le vrai modèle θ_0 , nous avons

$$Y_t = F(Y_{t-1}; \theta_0) + \eta_t.$$

Nous posons,

$$\Delta F_{t-1} = F(Y_{t-1}; \theta_0) - F(Y_{t-1}; \theta). \quad (3.16)$$

Ainsi, nous trouvons

$$\varphi_n(\theta) - \varphi_n(\theta_0) = \frac{A_n}{n} + \frac{B_n}{n}, \quad (3.17)$$

avec

$$A_n = \sum_{t=1}^n \|\Delta F_{t-1}\|_1^2 \quad \text{et} \quad B_n = 2 \sum_{t=1}^n (\eta_t \mid \Delta F_{t-1}), \quad (3.18)$$

où $(\cdot \mid \cdot)$ est le produit scalaire associé à la norme $\|\cdot\|_1$.

Notons que,

$$\Delta F_{t-1} = \begin{pmatrix} f(Y_{t-1}; \theta_0) - f(Y_{t-1}; \theta) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Il en résulte que,

$$A_n = \sum_{t=1}^n (f(Y_{t-1}; \theta) - f(Y_{t-1}; \theta_0))^2 \text{ et } B_n = 2 \sum_{t=1}^n \varepsilon_t (f(Y_{t-1}; \theta) - f(Y_{t-1}; \theta_0)). \quad (3.19)$$

Comme Θ est compact, il existe $B > 0$ tel que $|\lambda| \leq B$. Soit c une constante générique dont la valeur exacte peut changer pendant le développement mathématique.

Ainsi, pour tout $x \in E$, nous avons

$$\begin{aligned} |f(x; \theta)| &= \left| \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle \right| = \left| \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle \right| \\ &\leq |\alpha_1| |x_1| + \cdots + |\alpha_p| |x_p| + |\lambda| + \frac{1}{2} \\ &\leq |\alpha_1| |x_1| + \cdots + |\alpha_p| |x_p| + B + \frac{1}{2} \\ &\leq c (1 + \|x\|_1). \end{aligned}$$

Ce qui implique que

$$(f(x; \theta))^2 \leq c (1 + \|x\|_1)^2.$$

Par conséquent, comme $\mu_{\theta_0}(\|\cdot\|_1^k) < \infty$ où $k \geq 2$, nous obtenons

$$f(\cdot; \theta) \in L^1(\mu_{\theta_0}).$$

Ainsi, en utilisant le théorème ergodique sur la chaîne (Y_t) , nous trouvons

$$\frac{A_n}{n} \xrightarrow{p.s.} \mu_{\theta_0} \left((f(\cdot; \theta) - f(\cdot; \theta_0))^2 \right). \quad (3.20)$$

Nous posons,

$$M_n = \frac{B_n}{2} = \sum_{t=1}^n \varepsilon_t (f(Y_{t-1}; \theta) - f(Y_{t-1}; \theta_0)). \quad (3.21)$$

Il est simple de vérifier que (M_n) est une martingale de carré intégrable par rapport à \mathcal{F} . Son processus croissant, noté $([M]_n)$, est défini par

$$[M]_n = \sum_{t=1}^n (f(Y_{t-1}; \theta) - f(Y_{t-1}; \theta_0))^2 \mathbb{E}|\varepsilon_t|^2. \quad (3.22)$$

Ainsi, presque sûrement nous avons

$$\frac{[M]_n}{n} \longrightarrow \mu_{\theta_0} \left((f(\cdot; \theta) - f(\cdot; \theta_0))^2 \Gamma \right) \geq 0, \quad \text{où } \Gamma = \mathbb{E}|\varepsilon_t|^2 < \infty.$$

Il en découle,

$$\frac{M_n}{n} \xrightarrow{p.s.} 0.$$

Par suite,

$$\frac{B_n}{n} \xrightarrow{p.s.} 0.$$

– Finalement, notre but est de montrer que

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0.$$

Rappelons que, pour tout $\theta \in \Theta$, nous avons

$$\varphi_n(\theta) = \mathbb{P}_n g_\theta \text{ et } K(\theta) = \sigma_{\theta_0} g_\theta,$$

où \mathbb{P}_n est la mesure empirique engendrée par les observations $Z_1 = (Y_1, Y_0), \dots$,

$Z_n = (Y_n, Y_{n-1})$ et σ_{θ_0} est la mesure invariante de la chaîne (Z_t) .

Soit $\mathcal{G} = \{g_\theta, \theta \in \Theta\}$ la classe des fonctions de contraste, indexée par Θ compact.

D'après la preuve du Lemme 3.2.1, Pour tout $\theta \in \Theta$ et $z \in E^2$, nous avons

$$g_\theta(z) \leq A(z), \quad (3.23)$$

où

$$A(z) = c (1 + |z|^2), \quad (3.24)$$

avec c est une constante et $|\cdot|$ est la norme sur $(\mathbb{R}^p)^2$ associée à la norme $\|\cdot\|_1$ définie sur \mathbb{R}^p . De plus, nous avons $A \in L^1(\sigma_{\theta_0})$. Ainsi, la fonction A s'appelle la fonction enveloppe de la classe \mathcal{G} .

Soit $q > 0$ fixé. Ainsi, pour tout $\theta \in \Theta$, nous avons

$$\begin{aligned} |\varphi_n(\theta) - K(\theta)| &= |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| \\ &= |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{|z| < q} + (\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{|z| > q}| \\ &\leq |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{|z| < q}| + \mathbb{P}_n(|g_\theta| \mathbb{1}_{|z| > q}) + \sigma_{\theta_0}(|g_\theta| \mathbb{1}_{|z| > q}) \\ &\leq |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{|z| < q}| + \mathbb{P}_n(A(z) \mathbb{1}_{|z| > q}) + \sigma_{\theta_0}(A(z) \mathbb{1}_{|z| > q}). \end{aligned}$$

Nous posons,

$$p_z = \sigma_{\theta_0}(z) \text{ et } p_z^n = \mathbb{P}_n(z).$$

Alors,

$$\begin{aligned} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta \mathbb{1}_{|z| < q}| &= \left| \sum_{|z| < q} g_\theta(z)(p_z^n - p_z) \right| \\ &\leq \sum_{|z| < q} A(z) |p_z^n - p_z|. \end{aligned}$$

Donc, nous obtenons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| = \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_\theta| \leq \sum_{|z| < q} A(z) |p_z^n - p_z| + (\mathbb{P}_n + \sigma_{\theta_0})(A(z) \mathbb{1}_{|z| > q}).$$

Comme $p_z^n \rightarrow p_z$ presque sûrement quand $n \rightarrow \infty$, alors la somme finie

$$\sum_{|z| < q} A(z) |p_z^n - p_z| \xrightarrow{p.s.} 0.$$

Par conséquent,

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_{\theta}| \leq 0 + 2\sigma_{\theta_0}(A(z)\mathbb{1}_{|z| > q}), \text{ p.s.}$$

En prenant $q \nearrow \infty$, nous obtenons presque sûrement,

$$\limsup_n \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g_{\theta}| = 0.$$

□

Comme pour le modèle RINAR(1), l'identifiabilité du modèle RINAR(p) n'est pas standard à cause de l'opérateur d'arrondi. Dans les deux paragraphes suivants nous traitons ce problème. Nous constatons deux cas dépendants des coefficients de régression α_j^* . Notamment, si tous les coefficients α_j^* sont rationnels, il y a une manque d'identifiabilité autour de λ^* , la vraie valeur du paramètre λ .

Par la suite, nous définissons sur l'espace des paramètres Θ la distance d par

$$d(\theta, \theta') = \max \left\{ |\alpha_j - \alpha'_j|, 1 \leq j \leq p, |\lambda - \lambda'| \right\}, \quad (3.25)$$

pour tout $\theta = (\alpha_1, \dots, \alpha_p, \lambda)$, et $\theta' = (\alpha'_1, \dots, \alpha'_p, \lambda') \in \Theta$.

3.2.1 La consistance forte de l'estimateur des moindres carrés quand un au moins des coefficients de régression α_j^* est irrationnel

Dans cette section, nous supposons qu'il existe au moins $j \in \{1, \dots, p\}$ tel que $\alpha_j^* \in \mathbb{R} \setminus \mathbb{Q}$, et nous examinons le problème d'identifiabilité du modèle dans ce cas.

Proposition 3.2.2. *Supposons que [H2] soit vérifiée. S'il existe au moins $j \in \{1, \dots, p\}$ tel que α_j^* est irrationnel. Alors, pour tout $\theta \in \Theta$,*

$$f(x; \theta) = f(x; \theta_0), \forall x = (x_i) \in E = \mathbb{Z}^p \iff \theta = \theta_0.$$

Preuve. Notre objectif est de montrer que s'il existe au moins $j \in \{1, \dots, p\}$ tel que $\alpha_j^* \in \mathbb{R} \setminus \mathbb{Q}$, alors pour tout $\theta = (\alpha_1, \dots, \alpha_p, \lambda) \in \Theta$ nous avons

$$\left\langle \sum_{i=1}^p \alpha_i x_i + \lambda \right\rangle = \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda^* \right\rangle, \forall x = (x_i) \in E \implies \lambda = \lambda^* \text{ et } \alpha_i = \alpha_i^*, \forall i = 1, \dots, p.$$

L'idée de la preuve est de montrer que si $\theta \neq \theta_0$ alors il existe $y_0 \in E = \mathbb{Z}^p$ tel que

$f(y_0; \theta) \neq f(y_0; \theta_0)$. Tout d'abord, nous supposons que $\alpha_1 \neq \alpha_1^*$.

Ceci implique que $|\alpha_1 - \alpha_1^*| > 0$. Comme Θ est compact, il existe $B > 0$ tel que $|\lambda| \leq B$.

De plus, il existe $x_0 \in \mathbb{Z}$ tel que

$$|\alpha_1 - \alpha_1^*||x_0| - 2B - 1 > 0.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Nous avons,

$$\begin{aligned} |f(y_0; \theta) - f(y_0; \theta_0)| &= |\langle \alpha_1 x_0 + \lambda \rangle - \langle \alpha_1^* x_0 + \lambda^* \rangle| \\ &\geq |\alpha_1 - \alpha_1^*||x_0| - 2B - 1 > 0. \end{aligned}$$

Ainsi, s'il existe $i \in \{1, \dots, p\}$ tel que $\alpha_i \neq \alpha_i^*$, alors il existe $y_0 \in E$ tel que

$$|f(y_0; \theta) - f(y_0; \theta_0)| > 0.$$

De plus, si $y_0 = 0_E$, alors

$$f(y_0; \theta) = f(y_0; \theta_0) \implies \langle \lambda \rangle = \langle \lambda^* \rangle.$$

Par suite, sans utiliser l'hypothèse d'irrationalité de l'un des coefficients de régression, nous trouvons

$$f(x; \theta) = f(x; \theta_0), \forall x \in E \implies \langle \lambda \rangle = \langle \lambda^* \rangle \text{ et } \alpha_i = \alpha_i^* \text{ pour tout } i = 1, \dots, p.$$

Maintenant, pour simplifier les notations, nous supposons que $j = 1$. Ainsi, nous avons $\alpha_1^* \in \mathbb{R} \setminus \mathbb{Q}$. De plus, rappelons que $\{\cdot\}$ représente la partie fractionnaire.

Nous supposons, sans perte de généralité, $\alpha_1^* > 0$, $\lambda \geq \lambda^* \geq 0$. Nous distinguons trois cas.

– Supposons que $\{\lambda^*\} \in \left[0, \frac{1}{2}\right]$. Il en découle,

$$\{\lambda\} \in \left[0, \frac{1}{2}\right] \text{ et } \lambda - \lambda^* = \{\lambda\} - \{\lambda^*\} \geq 0.$$

Comme $\alpha_1^* \in \mathbb{R} \setminus \mathbb{Q}$, alors $(\{\alpha_1^* x\})_{x \in \mathbb{Z}}$ est dense dans $[0, 1[$. Ainsi, si $\{\lambda\} \neq \{\lambda^*\}$, alors il existe $x_0 \in \mathbb{Z}$ tel que

$$\{\alpha_1^* x_0\} + \{\lambda\} \geq \frac{1}{2} \text{ et } \{\alpha_1^* x_0\} + \{\lambda^*\} < \frac{1}{2}.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Nous trouvons

$$\begin{aligned} f(y_0; \theta) - f(y_0; \theta_0) &= \langle \alpha_1^* x_0 + \lambda \rangle - \langle \alpha_1^* x_0 + \lambda^* \rangle \\ &= \langle \{\alpha_1^* x_0\} + \{\lambda\} \rangle - \langle \{\alpha_1^* x_0\} + \{\lambda^*\} \rangle \\ &= 1. \end{aligned}$$

– Supposons que $\{\lambda^*\} \in \left[\frac{1}{2}, 1\right[$ et $\{\lambda\} \in \left[\frac{1}{2}, 1\right[$. Il en résulte,

$$\lambda - \lambda^* = \{\lambda\} - \{\lambda^*\} \geq 0.$$

Ainsi, si $\{\lambda\} \neq \{\lambda^*\}$, alors il existe $x_0 \in \mathbb{Z}$ tel que

$$\{\alpha_1^* x_0\} + \{\lambda\} \geq \frac{3}{2} \text{ et } \{\alpha_1^* x_0\} + \{\lambda^*\} < \frac{3}{2}.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Nous obtenons

$$\begin{aligned} f(y_0; \theta) - f(y_0; \theta_0) &= \langle \{\alpha_1^* x_0\} + \{\lambda\} \rangle - \langle \{\alpha_1^* x_0\} + \{\lambda^*\} \rangle \\ &= 1. \end{aligned}$$

– Supposons que $\{\lambda^*\} \in \left[\frac{1}{2}, 1\right[$ et $\{\lambda\} \in \left[0, \frac{1}{2}\right[$. Ainsi,

$$[\lambda] = [\lambda^*] + 1 \text{ et } \lambda - \lambda^* = 1 + \{\lambda\} - \{\lambda^*\} \geq 0.$$

Ainsi, il existe $x_0 \in \mathbb{Z}$ tel que

$$\{\alpha_1^* x_0\} + \{\lambda\} \geq \frac{1}{2} \text{ et } \{\alpha_1^* x_0\} + \{\lambda^*\} < \frac{3}{2}.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Nous avons

$$\begin{aligned} f(y_0; \theta) - f(y_0; \theta_0) &= 1 + \langle \{\alpha_1^* x_0\} + \{\lambda\} \rangle - \langle \{\alpha_1^* x_0\} + \{\lambda^*\} \rangle \\ &= 1. \end{aligned}$$

□

Proposition 3.2.3. *Supposons que [H2] soit vérifiée. S'il existe au moins $j \in \{1, \dots, p\}$ tel que α_j^* est irrationnel. Alors, pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_\varepsilon} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}$.

Avant de montrer cette proposition, nous donnons quelques rappels nécessaires.

Supposons que [H2] soit vérifiée, alors d'après la Proposition 3.2.1 nous avons,

$$K(\theta) - K(\theta_0) = \sigma_{\theta_0} [g_\theta(\cdot) - g_{\theta_0}(\cdot)] = \mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right). \quad (3.26)$$

Soit $x_0 \in \mathbb{N}^*$ fixé, la fonction $\alpha \longrightarrow \{\alpha x_0\}$, définie sur \mathbb{R} et à valeurs dans $[0, 1[$, est continue en α_0 si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$.

Preuve. Rappelons que les composantes du paramètre θ ne possèdent pas les mêmes rôles.

Pour cela, nous distinguons trois situations pour l'événement Θ_ε .

Ainsi, notons que $\Theta_\varepsilon = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, où

$$\Gamma_1 = \{\theta \in \Theta, \exists i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| \geq \varepsilon\}, \quad (3.27)$$

$$\Gamma_2 = \{\theta \in \Theta, \forall i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| < \varepsilon, |\lambda - \lambda^*| \geq \varepsilon, \langle \lambda \rangle \neq \langle \lambda^* \rangle\}, \quad (3.28)$$

et

$$\Gamma_3 = \{\theta \in \Theta, \forall i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| < \varepsilon, |\lambda - \lambda^*| \geq \varepsilon, \langle \lambda \rangle = \langle \lambda^* \rangle\}. \quad (3.29)$$

Nous allons montrer

$$\inf_{\theta \in \Gamma_i} |K(\theta) - K(\theta_0)| > 0, \quad i = 1, 2, 3.$$

L'idée de la preuve est basée sur l'équation (3.26). Alors, l'objectif est de trouver $y_0 \in E$ tel que

$$|f(y_0; \theta) - f(y_0; \theta_0)| > 0, \quad \text{uniformément sur } \Gamma_i, \quad i = 1, 2, 3.$$

- Nous considérons le premier cas, $\theta \in \Gamma_1$. Pour simplifier les notations, nous supposons que $i = 1$. Ainsi, nous avons $|\alpha_1 - \alpha_1^*| \geq \varepsilon > 0$. Comme Θ est compact, il existe une constante $B > 0$ tel que $|\lambda| \leq B$. De plus, il existe $x_0 \in \mathbb{Z}^*$ tel que

$$\varepsilon|x_0| - 2B - 1 > 0.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Nous avons

$$\begin{aligned} |f(y_0; \theta) - f(y_0; \theta_0)| &\geq \varepsilon|x_0| - 2B - 1 \\ &> 0. \end{aligned}$$

D'ailleurs pour tout $y = (x, 0, \dots, 0) \in E$ tel que $|x| \geq |x_0|$ nous avons

$\mu_{\theta_0} \left([f(\cdot; \theta) - f(\cdot; \theta_0)]^2 \right) > 0$ uniformément sur Γ_1 . Il en découle, puisque le support de μ_{θ_0} n'est pas borné,

$$\inf_{\theta \in \Gamma_1} |K(\theta) - K(\theta_0)| > 0.$$

- Supposons que $\theta \in \Gamma_2$. Ici, nous avons $\langle \lambda \rangle \neq \langle \lambda^* \rangle$. De plus, la valeur absolue de la soustraction de deux entiers différents est supérieur ou égale à 1, c.à.d.

$$|\langle \lambda \rangle - \langle \lambda^* \rangle| \geq 1.$$

Soit $y_0 = 0_E = (0, 0, \dots, 0)$. Ainsi, nous trouvons

$$|f(y_0; \theta) - f(y_0; \theta_0)| = |\langle \lambda \rangle - \langle \lambda^* \rangle| \geq 1.$$

Donc,

$$\inf_{\theta \in \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

- Finalement, nous considérons le dernier cas, $\theta \in \Gamma_3$. Pour simplifier les notations nous supposons, sans perte de généralité, $\alpha_1^* \in \mathbb{R} \setminus \mathbb{Q}$. Soient

$$E' = \{x = (x_1, \dots, x_p)^\tau \in E, x_1 \in \mathbb{Z} \text{ et } x_i = 0, \forall i = 2, \dots, p\}$$

et

$$\Lambda = \{\theta \in \Theta, |\alpha_1 - \alpha_1^*| < \varepsilon, |\lambda - \lambda^*| \geq \varepsilon, \langle \lambda \rangle = \langle \lambda^* \rangle\}.$$

D'après la preuve de la Proposition 2.4.3, pour tout $\theta \in \Lambda$, il existe $x_0 \in \mathbb{Z}$ tel que

$$|\langle \alpha_1 x_0 + \lambda \rangle - \langle \alpha_1^* x_0 + \lambda^* \rangle| > 0.$$

Sur E' , le modèle RINAR(p) actuel devient un processus RINAR(1) avec le coefficient de régression $\alpha_1^* \in \mathbb{R}$ irrationnel.

Ainsi, pour tout $\theta \in \Gamma_3$, il existe $y_0 = (x_0, 0, \dots, 0)^\tau \in E' \subset E$ tel que

$$|f(y_0; \theta) - f(y_0; \theta_0)| = |\langle \alpha_1 x_0 + \lambda \rangle - \langle \alpha_1^* x_0 + \lambda^* \rangle| > 0.$$

Par suite,

$$\inf_{\theta \in \Gamma_3} |K(\theta) - K(\theta_0)| > 0.$$

□

Théorème 3.2.1. *Supposons que [H2] soit vérifiée. S'il existe au moins $j \in \{1, \dots, p\}$ tel que α_j^* est irrationnel. Alors, l'estimateur des moindres carrés est fortement consistant, c.à.d.*

$$\hat{\alpha}_n \rightarrow \alpha_0, \quad \mathbb{P}_{\alpha_0} - p.s.$$

Preuve. La conclusion $\hat{\theta}_n \rightarrow \theta_0$ presque sûrement résulte d'après les Propositions 3.2.1 et 3.2.3, par des arguments standards de la théorie des M-estimateurs (voir par exemple Van Der Vaart (1998, [61]), Théorème 5.7). □

3.2.2 La consistance forte de l'estimateur des moindres carrés quand tous coefficients de régression α_j^* sont rationnels

Dans cette section, nous supposons que tous les coefficients de régression, sous le vrai modèle, sont rationnels. Ceci implique que, pour tout $j \in \{1, \dots, p\}$, nous avons

$$\alpha_j^* = \frac{a_j}{b_j}, \text{ où } a_j \in \mathbb{Z} \text{ } b_j \in \mathbb{N}^* \text{ et } a_j \wedge b_j = 1.$$

Soient $x = (x_1, \dots, x_p)^T \in E = \mathbb{Z}^p$ et $\alpha_0 = (\alpha_1^*, \dots, \alpha_p^*)$. Dans la suite, $\langle \cdot | \cdot \rangle$ représente le produit scalaire usuel. Ainsi, nous avons

$$\langle \alpha_0 | x \rangle = \sum_{i=1}^p \alpha_i^* x_i = \frac{1}{\prod_{j=1}^p b_j} \left(\sum_{l=1}^p A_l x_l \right), \quad (3.30)$$

où

$$A_l = a_l \prod_{j=1, j \neq l}^p b_j. \quad (3.31)$$

D'après le théorème de Bézout, nous obtenons

$$A_1 \mathbb{Z} + \dots + A_p \mathbb{Z} = d \mathbb{Z}, \quad (3.32)$$

où $d = A_1 \wedge \dots \wedge A_p$ est le P.G.C.D. de A_1, \dots, A_p . Alors, il existe $y \in \mathbb{Z}$ tel que

$$\sum_{i=1}^p \alpha_i^* x_i = \nu_0 y, \quad (3.33)$$

avec

$$\nu_0 = \frac{d}{\prod_{j=1}^p b_j} \in]-1, 1[. \quad (3.34)$$

Notons que le numérateur et le dénominateur de ν_0 ne sont pas nécessairement premier entre eux. Ainsi, nous réécrivons ν_0 avec sa forme de fraction irréductible

$$\nu_0 = \frac{a}{b}, \quad (3.35)$$

où $a \in \mathbb{Z}$, $b \in \mathbb{N}^*$ et $a \wedge b = 1$.

Proposition 3.2.4. *Supposons que [H2] soit vérifiée. Soit $\theta = (\alpha_1^*, \dots, \alpha_p^*, \lambda) \in \Theta$.*

Si pour tout $j \in \{1, \dots, p\}$, nous avons $\alpha_j^ = \frac{a_j}{b_j}$, où $a_j \in \mathbb{Z}$, $b_j \in \mathbb{N}^*$ et $a_j \wedge b_j = 1$.*

Alors,

$$f(y; \theta) = f(y; \theta_0), \quad \forall y \in E \iff \langle \nu_0 x + \lambda \rangle = \langle \nu_0 x + \lambda^* \rangle, \quad \forall x \in \mathbb{Z}.$$

Preuve. Soit $y = (x_1, \dots, x_p)^T \in E$. Alors, d'après (3.33) (3.34) et (3.35), il existe $x_0 \in \mathbb{Z}$ tel que $\sum_{i=1}^p \alpha_i^* x_i = \nu_0 x_0$. Par suite,

$$\begin{aligned} f(y; \theta) = f(y; \theta_0) &\implies \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda \right\rangle = \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda^* \right\rangle \\ &\implies \langle \nu_0 x_0 + \lambda \rangle = \langle \nu_0 x_0 + \lambda^* \rangle. \end{aligned}$$

D'autre part, soit $x \in \mathbb{Z}$. D'après (3.34), nous avons $\nu_0 x = \frac{d}{\prod_{j=1}^p b_j} x$.

De plus, d'après (3.32), il existe $y_0 = (x_1, \dots, x_p)^T \in E$ tel que

$$d x = A_1 x_1 + \dots + A_p x_1.$$

Il en résulte

$$\nu_0 x = \frac{1}{\prod_{j=1}^p b_j} (A_1 x_1 + \cdots + A_p x_1) = \sum_{i=1}^p \alpha_i^* x_i = \langle \alpha_0 \mid y_0 \rangle.$$

Ainsi,

$$\begin{aligned} \langle \nu_0 x + \lambda \rangle = \langle \nu_0 x + \lambda^* \rangle &\implies \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda \right\rangle = \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda^* \right\rangle \\ &\implies f(y_0; \theta) = f(y_0; \theta_0). \end{aligned}$$

□

Proposition 3.2.5. *Supposons que [H2] soit vérifiée et pour tout $j \in \{1, \dots, p\}$, nous avons $\alpha_j^* = \frac{a_j}{b_j}$, où $a_j \in \mathbb{Z}$, $b_j \in \mathbb{N}^*$ et $a_j \wedge b_j = 1$. Alors, pour tout $\theta \in \Theta$,*

$$f(y; \theta) = f(y; \theta_0) \quad \forall y \in E \iff \alpha_j = \alpha_j^*, \quad \forall j \in \{1, \dots, p\}, \quad \text{et } \lambda \in I_0,$$

où $I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda^* \rangle\}$. De plus, ce dernier est de taille $\frac{1}{b}$, où b est le dénominateur de ν_0 défini par (3.34) et (3.35).

Avant de montrer la proposition ci-dessus, quelques précisions sont nécessaires.

La localisation de I_0 dépend de la parité de b et de la position de la partie fractionnaire de λ^* (c.à.d. $\{\lambda^*\}$) sur l'intervalle $[0, 1[$. Ainsi, nous distinguons deux cas :

- Supposons que b est pair. Alors, I_0 est défini par (2.65) et (2.66).
- Supposons que b est impair. Nous remarquons trois sous-cas :

- Si $\{\lambda_0\} \in \left[0, \frac{1}{2b}\right]$, alors I_0 est défini par (2.67).
- Si $\{\lambda_0\} \in \left[\frac{1}{2b}, \frac{2b-1}{2b}\right]$, alors I_0 est défini par (2.68) et (2.69).
- Si $\{\lambda_0\} \in \left[\frac{2b-1}{2b}, 1\right]$, alors I_0 est défini par (2.70).

Notons que, pour rester compatible avec les notations du chapitre présent, il suffit de remplacer q par b et λ_0 par λ^* dans les équations précédentes.

Maintenant, nous détaillons le cas où b est pair. Ainsi, il existe $k_0 \in \{0, 1, \dots, q-1\}$

tel que

$$\{\lambda^*\} \in i_0 = \left[\frac{k_0}{b}, \frac{k_0+1}{b}\right] \quad (3.36)$$

Par suite, nous définissons

$$I_0 = \{\lambda, \langle \lambda \rangle = \langle \lambda^* \rangle \text{ et } \{\lambda\} \in i_0\}. \quad (3.37)$$

Ici, nous donnons un exemple numérique. Nous considérons le modèle RINAR(4) avec

$$\alpha_0 = (\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_4^*) = \left(\frac{3}{25}, \frac{3}{8}, \frac{1}{5}, -\frac{1}{4} \right) \quad \text{et } \lambda^* = 2.5.$$

D'après (3.31), nous obtenons

$$\prod_{j=1}^4 b_j = 800, \quad A_1 = 480, \quad A_2 = 1500, \quad A_3 = 800 \quad \text{et } A_4 = -1000.$$

Alors, nous avons $d = A_1 \wedge A_2 \wedge A_3 \wedge A_4 = 20$. Donc, d'après (3.34), nous trouvons

$$\nu_0 = \frac{20}{800}.$$

La forme irréductible de ν_0 est

$$\nu_0 = \frac{a}{b} = \frac{1}{40} = 0.025.$$

Par conséquent, $b = 40$ est pair. De plus, nous avons $\{\lambda^*\} = 0.5$ et $\langle \lambda^* \rangle = 3$.

Donc, en utilisant la subdivision de l'intervalle $[0, 1[$ en 40 intervalles de taille $\frac{1}{40}$, nous trouvons

$$\{\lambda^*\} \in i_0 = [0.5, 0.525[$$

et

$$I_0 = [2.5, 2.525[.$$

Preuve. D'après la preuve de la Proposition 3.2.2, pour tout $\theta \in \Theta$, nous avons

$$f(y; \theta) = f(y; \theta_0), \quad \forall y \in E \Rightarrow \langle \lambda \rangle = \langle \lambda^* \rangle \text{ et } \alpha_i = \alpha_i^* \text{ pour tout } i = 1, \dots, p.$$

Soit

$$\Xi = \{\theta \in \Theta, \alpha_i = \alpha_i^* \quad \forall i = 1, \dots, p \text{ et } \lambda \in I\}, \quad \text{où } I = \{\lambda, \langle \lambda \rangle = \langle \lambda^* \rangle\}.$$

Ainsi, d'après la Proposition 3.2.4, pour tout $\theta \in \Xi$, nous avons

$$f(y; \theta) = f(y; \theta_0), \quad \forall y \in E \iff \langle \nu_0 x + \lambda \rangle = \langle \nu_0 x + \lambda^* \rangle, \quad \forall x \in \mathbb{Z}.$$

Par suite, d'après la preuve de la Proposition 2.4.4, nous déduisons que I est de taille $\frac{1}{b}$ ($I \equiv I_0$). D'autre part, soit

$$\Xi_0 = \{\theta \in \Theta, \alpha_i = \alpha_i^* \quad \forall i = 1, \dots, p \text{ et } \lambda \in I_0\}.$$

Rappelons que, d'après le théorème de Bézout, pour tout $y = (x_1, \dots, x_p)^\tau \in E$

il existe $x \in \mathbb{Z}$ tel que $\sum_{i=1}^p \alpha_i^* x_i = \nu_0 x$. Soit $\theta \in \Xi_0$.

Ainsi, $\forall y \in E$ nous avons

$$\begin{aligned} |f(y; \theta) - f(y; \theta_0)| &= \left| \left\langle \sum_{i=1}^p \alpha_i^* x_i + \lambda \right\rangle - \left\langle \alpha_i^* x_i + \lambda^* \right\rangle \right| \\ &= \left| \langle \nu_0 x + \lambda \rangle - \langle \nu_0 x + \lambda^* \rangle \right| \\ &= 0. \end{aligned}$$

□

Soient $\alpha_0 = (\alpha_1^*, \dots, \alpha_p^*)$ et $y_0 = (x_1, \dots, x_p)^\tau \in E$. Rappelons que, d'après le théorème de Bézout, il existe $x_0 \in \mathbb{Z}$ tel que

$$\langle \alpha_0 \mid y_0 \rangle = \sum_{j=1}^p \alpha_j^* x_j = \nu_0 x_0 = \frac{a}{b} x_0.$$

Lemme 3.2.2. *La fonction $\alpha = (\alpha_1, \dots, \alpha_p) \longrightarrow \left\{ \sum_{j=1}^p \alpha_j x_j^0 \right\}$, définie sur \mathbb{R}^p et à valeurs dans $[0, 1[$, est continue en α_0 si $x_0 \neq mb$, où $m \in \mathbb{N}$ (x_0 n'est pas un multiple de b)).*

Preuve. Soit $y_0 = (x_1, \dots, x_p)^\tau \in E$ fixé. La fonction $\alpha \longrightarrow \{\langle \alpha \mid y_0 \rangle\}$, définie sur \mathbb{R}^p , est discontinue en α_0 si $\langle \alpha_0 \mid y_0 \rangle \in \mathbb{Z}$. D'autre part, nous avons

$$\langle \alpha_0 \mid y_0 \rangle = \frac{a}{b} x_0.$$

Il en résulte, comme $x_0 \neq mb$, $\langle \alpha_0 \mid y_0 \rangle \notin \mathbb{Z}$.

□

Quelques rappels et notations sont nécessaires. Nous posons,

$$E_0 = \{\theta \in \Theta, \alpha_j = \alpha_j^* \forall j = 1, \dots, p \text{ et } \lambda \in I_0\} = \{\alpha_0\} \times I_0 \text{ où } \alpha_0 = (\alpha_1^*, \dots, \alpha_p^*).$$

Notons que $\theta_0 = (\alpha_1^*, \dots, \alpha_p^*, \lambda^*) \in E_0$. Sous [H2] et d'après la Proposition 3.2.1 nous avons,

$$K(\theta) - K(\theta_0) = \mu_{\theta_0} [f(\cdot; \theta) - f(\cdot; \theta_0)]^2.$$

D'autre part, d est la distance sur Θ est définie par

$$d(\theta, \theta') = \max \left\{ |\alpha_1 - \alpha'_1|, \dots, |\alpha_p - \alpha'_p|, |\lambda - \lambda'| \right\},$$

pour tout $\theta = (\alpha_1, \dots, \alpha_p, \lambda), \theta' = (\alpha'_1, \dots, \alpha'_p, \lambda') \in \Theta$.

Soit $\theta_* \in E_0$. D'après la Proposition 3.2.5, nous avons

$$K(\theta_*) - K(\theta_0) = \mu_{\theta_0} [f(\cdot; \theta_*) - f(\cdot; \theta_0)]^2 = 0.$$

Proposition 3.2.6. *Supposons que [H2] soit vérifiée et pour tout $j \in \{1, \dots, p\}$, nous avons $\alpha_j^* = \frac{a_j}{b_j}$, où $a_j \in \mathbb{Z}$, $b_j \in \mathbb{N}^*$ et $a_j \wedge b_j = 1$. Alors, pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_\varepsilon^0} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon^0 = \{\theta : d(\theta, E_0) \geq \varepsilon\}$.

Preuve. Pour simplifier la présentation, nous supposons que b , le dénominateur de ν_0 , est pair. Ainsi, il existe $k_0 \in \{0, 1, \dots, q-1\}$ tel que

$$\{\lambda^*\} \in i_0 = \left[\frac{k_0}{b}, \frac{k_0+1}{b} \right[.$$

Par suite, nous avons

$$I_0 = \{\lambda : \langle \lambda \rangle = \langle \lambda^* \rangle, \{\lambda\} \in i_0\} \text{ et } E_0 = \{\alpha_0\} \times I_0.$$

Nous distinguons trois situations différentes pour l'événement $\Theta_\varepsilon^0 = \{\theta : d(\theta, E_0) \geq \varepsilon\}$. Ainsi, nous avons $\Theta_\varepsilon^0 = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, avec

$$\Gamma_1 = \{\theta : \exists i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| \geq \varepsilon\}, \quad (3.38)$$

$$\Gamma_2 = \{\theta : \forall i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| < \varepsilon, d(\lambda, I_0) \geq \varepsilon, \langle \lambda \rangle \neq \langle \lambda^* \rangle\}, \quad (3.39)$$

et

$$\Gamma_3 = \{\theta : \forall i \in \{1, \dots, p\}, |\alpha_i - \alpha_i^*| < \varepsilon, d(\lambda, I_0) \geq \varepsilon, \langle \lambda \rangle = \langle \lambda^* \rangle\}. \quad (3.40)$$

Notons que

$$d(\lambda, I_0) = \inf_{\lambda_* \in I_0} |\lambda - \lambda_*| \leq |\lambda - \lambda^*|.$$

Nous allons montrer

$$\inf_{\theta \in \Gamma_i} |K(\theta) - K(\theta_0)| > 0, i = 1, 2, 3.$$

L'idée de la preuve est de trouver $y_0 \in E$ tel que

$$|f(y_0; \theta) - f(y_0; \theta_*)| > 0, \text{ uniformément sur } \Gamma_i, i = 1, 2, 3.$$

– Pour $\theta \in \Gamma_1 \cup \Gamma_2$, d'après la preuve de la Proposition 3.2.3, nous avons

$$\inf_{\theta \in \Gamma_1 \cup \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

– Nous considérons le cas où $\theta \in \Gamma_3$. Ici, pour simplifier les notations, nous supposons sans perte de généralité $\lambda^* \geq 0$ (par conséquent $I_0 \subset \mathbb{R}^+$), et $\nu_0 = \frac{1}{4}$ (c.à.d. $b = 4$). Nous distinguons quatre cas dépendants de la position de la partie fractionnaire de λ^* et λ :

1. Cas A : $\{\lambda^*\} \in [0, \frac{1}{2}[$ et $\lambda \geq \lambda_0$.
 2. Cas B : $\{\lambda^*\} \in [\frac{1}{2}, 1[$ et $\lambda \geq \lambda_0$.
 3. Cas C : $\{\lambda^*\} \in [0, \frac{1}{2}[$ et $\lambda \leq \lambda_0$.
 4. Cas D : $\{\lambda^*\} \in [\frac{1}{2}, 1[$ et $\lambda \leq \lambda_0$.
- Cas A : nous avons

$$\{\lambda^*\} \in i_0 = \left[0, \frac{1}{4}\right[\text{ et } \{\lambda\} \in \left[\frac{1}{4}, \frac{1}{2}\right[.$$

Par conséquent,

$$d(\theta, E_0) = d(\lambda, I_0) = \inf_{\lambda_* \in I_0} |\lambda - \lambda_*| = \inf_{\{\lambda_*\} \in i_0} \{\lambda\} - \{\lambda_*\}.$$

Soit $x_0 = 1$ fixé. D'après le théorème de Bézout, il existe $y_0 = (x_1, \dots, x_p)^\tau \in E$ tel que

$$\sum_{i=1}^p \alpha_i^* x_i = \nu_0 x_0 = \frac{1}{4}.$$

Ainsi, pour tout $\{\lambda_*\} \in i_0$ (en particulier pour $\{\lambda^*\}$), nous avons

$$\left\{ \sum_{i=1}^p \alpha_i^* x_i \right\} + \{\lambda_*\} = \{\nu_0 x_0\} + \{\lambda_*\} = \left\{ \frac{1}{4} x_0 \right\} + \{\lambda_*\} = \frac{1}{4} + \{\lambda_*\} < \frac{1}{2} \quad (3.41)$$

et

$$\left\{ \sum_{i=1}^p \alpha_i^* x_i \right\} + \{\lambda\} = \{\nu_0 x_0\} + \{\lambda\} = \frac{1}{4} + \{\lambda\} \geq \frac{1}{2} + \varepsilon. \quad (3.42)$$

Comme $x_0 = 1$ n'est pas un multiple de $b = 4$, d'après le Lemme 3.2.2, la fonction $\alpha \rightarrow \{\sum_{i=1}^p \alpha_i^* x_i\}$ est continue en $\alpha_0 = (\alpha_1^*, \dots, \alpha_p^*)$. Ainsi, il existe $\eta = \eta(\varepsilon, \nu_0, \lambda^*) \leq \varepsilon$ tel que $\forall j = 1, \dots, p$ $|\alpha_j - \alpha_j^*| \leq \eta$ nous avons

$$\begin{cases} [\sum_{j=1}^p \alpha_j x_j] = [\sum_{j=1}^p \alpha_j^* x_j]. \\ |\left\{ \sum_{j=1}^p \alpha_j x_j \right\} - \left\{ \sum_{j=1}^p \alpha_j^* x_j \right\}| \leq \varepsilon. \end{cases} \quad (3.43)$$

Donc, d'après (3.42) et (3.43), nous obtenons

$$\begin{aligned} \left\{ \sum_{j=1}^p \alpha_j x_j \right\} + \{\lambda\} &= \left(\left\{ \sum_{j=1}^p \alpha_j x_j \right\} - \left\{ \sum_{j=1}^p \alpha_j^* x_j \right\} \right) + \left(\left\{ \sum_{j=1}^p \alpha_j^* x_j \right\} + \{\lambda\} \right) \\ &\geq -\varepsilon + \left(\frac{1}{2} + \varepsilon \right) = \frac{1}{2}. \end{aligned}$$

Ainsi, nous trouvons

$$\begin{aligned}
f(y_0; \theta) - f(y_0; \theta_0) &= \left\langle \sum_{j=1}^p \alpha_j x_j + \lambda \right\rangle - \left\langle \sum_{j=1}^p \alpha_j^* x_j + \lambda_* \right\rangle \\
&= \left[\sum_{j=1}^p \alpha_j x_j \right] - \left[\sum_{j=1}^p \alpha_j^* x_j \right] + \langle \lambda \rangle - \langle \lambda_* \rangle \\
&+ \left\langle \left\{ \sum_{j=1}^p \alpha_j x_j \right\} + \{\lambda\} \right\rangle - \left\langle \left\{ \sum_{j=1}^p \alpha_j^* x_j \right\} + \{\lambda_*\} \right\rangle \\
&= 1.
\end{aligned}$$

Les autres cas peuvent être traités de façon similaire, avec un choix convenable de x_0 et $\alpha = (\alpha_1, \dots, \alpha_p)$ suffisamment proche de $\alpha_0 = (\alpha_1^*, \dots, \alpha_p^*)$. Alors, nous déduisons qu'il existe $\eta = \eta(\varepsilon, \nu_0, \lambda^*) \leq \varepsilon$ et une constante $e_0 > 0$ tels que uniformément sur $\Gamma'_3 = \left\{ \theta \in \Gamma_3, |\alpha_j - \alpha_j^*| \leq \eta, \forall j = 1, \dots, p \right\}$, nous avons

$$|K(\theta) - K(\theta_0)| \geq e_0 > 0.$$

Pour $\theta \in \Gamma''_3 = \Gamma_3 \setminus \Gamma'_3 = \{ \theta \in \Gamma_3, \exists i \in \{1, \dots, p\} \text{ tel que } |\alpha_i - \alpha_i^*| > \eta \}$, les mêmes arguments pour Γ_1 peuvent être appliqués ici et par suite il existe $d_0 > 0$ tel que uniformément sur Γ''_3 , nous trouvons

$$|K(\theta) - K(\theta_0)| \geq d_0 > 0.$$

□

Théorème 3.2.2. *Supposons que [H2] soit vérifiée et pour tout $j \in \{1, \dots, p\}$, nous avons*

$\alpha_j^ = \frac{a_j}{b_j}$, où $a_j \in \mathbb{Z}$ $b_j \in \mathbb{N}^*$ et $a_j \wedge b_j = 1$. Alors,*

$$d(\hat{\theta}_n, E_0) \xrightarrow{p.s.} 0, \text{ où } E_0 = \{\alpha_0\} \times I_0.$$

En d'autres termes, $\hat{\alpha}_n$ est fortement consistant tandis que $\hat{\lambda}_n$ converge vers un intervalle de taille $\frac{1}{b}$ contenant λ^ .*

Preuve. La conclusion $d(\hat{\theta}_n, E_0) \longrightarrow 0$, presque sûrement, résulte des Propositions 3.2.1 et 3.2.6 par les mêmes arguments utilisés dans la preuve du Théorème 3.2.1. □

Soit A un intervalle dans \mathbb{R} , nous notons par $|A|$ la taille de A .

Dans le chapitre 2, nous avons vu que pour un modèle RINAR(1) avec un coefficient de régression $\alpha_0 = \frac{a_0}{b_0}$, où $a_0 \in \mathbb{Z}$ $b_0 \in \mathbb{N}^*$ et $a_0 \wedge b_0 = 1$, la taille de l'intervalle I_0 est $\frac{1}{b_0}$.

Ainsi, si $\alpha_0 = \frac{1}{2}$ l'intervalle I_0 atteint la taille maximale ($|I_0| = \frac{1}{2}$).

Ici, une question naturelle se pose :

Pouvons-nous réduire, avec un modèle d'ordre supérieur ($p > 1$) la taille de l'intervalle I_0 ?

Pour répondre à cette question, nous considérons le modèle RINAR(2) avec

$$\alpha_1^* = \frac{a_1}{b_1} \quad \text{et} \quad \alpha_2^* = \frac{a_2}{b_2}.$$

Ainsi, nous avons

$$\nu_0 = \frac{d}{b_1 b_2} = \frac{a}{b} \quad \text{où} \quad d = a_1 b_2 \wedge a_2 b_1 \quad \text{et} \quad a \wedge b = 1.$$

Par conséquent, nous obtenons

$$|I_0| = \frac{1}{b}.$$

Comme b divise $a(b_1 b_2)$ et $a \wedge b = 1$, alors b divise $b_1 b_2$ ($b \leq b_1 b_2$). Il en découle

$$\frac{1}{b_1 b_2} \leq \frac{1}{b} = |I_0|.$$

Sur $E' = \{(x, y) \in \mathbb{Z}^2, y = 0\}$, le modèle RINAR(2) présent devient un processus RINAR(1) avec α_1^* comme coefficient de régression. Par conséquent, nous obtenons un ensemble $I_{0,1}$ tel que $|I_{0,1}| = \frac{1}{b_1}$. En revanche, sur $E'' = \{(x, y) \in \mathbb{Z}^2, x = 0\}$, RINAR(2) devient un processus RINAR(1) avec α_2^* comme coefficient de régression. Par conséquent, nous obtenons un ensemble $I_{0,2}$ tel que $|I_{0,2}| = \frac{1}{b_2}$. Ainsi,

$$I_0 \subseteq I_{0,1} \cap I_{0,2}.$$

Ceci implique que

$$|I_0| \leq |I_{0,1} \cap I_{0,2}| \leq \min\left\{\frac{1}{b_1}, \frac{1}{b_2}\right\}.$$

Finalement,

$$\frac{1}{b_1 b_2} \leq |I_0| \leq \min\left\{\frac{1}{b_1}, \frac{1}{b_2}\right\}.$$

Notons que, sous l'hypothèse $|\alpha_1^*| + |\alpha_2^*| < 1$, si la valeur absolue d'un des deux coefficients est $\frac{1}{2}$, alors l'autre est strictement inférieur à $\frac{1}{2}$. Dans ce cas, la taille maximale de l'intervalle I_0 est strictement inférieur à $\frac{1}{2}$.

Par conséquent, en augmentant l'ordre de l'autorégression p , nous pouvons effectivement nous attendre à une réduction de l'intervalle I_0 où le paramètre λ reste identifiable.

3.3 Méthode numérique pour calculer $\hat{\theta}_n$

Pour calculer l'estimateur des moindres carrés, nous généralisons l'algorithme proposé dans le chapitre 2. Ainsi, l'algorithme généralisé possède deux étapes. La première consiste à préciser un point de départ significatif. Puis, l'algorithme continue par des étapes de recherche dichotomique successives.

3.3.1 Initialisation

Comme le processus RINAR(p) est inspiré du modèle AR(p) réel, nous proposons pour le point de départ de l'algorithme $\hat{\theta}_0 = (\hat{\alpha}_{1,0}, \dots, \hat{\alpha}_{p,0}, \hat{\lambda}_0)$, l'estimateur de Yule-Walker du paramètre $\theta = (\alpha_1, \dots, \alpha_p, \lambda)$. Ainsi, nous avons

$$\begin{pmatrix} \hat{\alpha}_{1,0} \\ \vdots \\ \hat{\alpha}_{p,0} \end{pmatrix} = \hat{R}_p^{-1} \hat{\rho}_p \quad \text{et} \quad \hat{\lambda}_0 = \bar{X}_n \left(1 - \sum_{j=1}^p \hat{\alpha}_{j,0}\right),$$

où $\hat{R}_p = [\hat{\rho}(i-j)]_{i,j=1}^p$ est la matrice d'autocorrélation empirique avec $\hat{\rho}_p = (\hat{\rho}(1), \dots, \hat{\rho}(p))^T$ et \bar{X}_n est la moyenne empirique.

3.3.2 Recherche dichotomique successive

Pour simplifier les notations, nous posons $p = 2$. Ainsi, le passage de $(\hat{\alpha}_{1,k}, \hat{\alpha}_{2,k}, \hat{\lambda}_k)$ à $(\hat{\alpha}_{1,k+1}, \hat{\alpha}_{2,k+1}, \hat{\lambda}_{k+1})$ se fait en $p + 1 = 3$ phases.

Dans la première phase nous fixons $\hat{\alpha}_{2,k}$ et $\hat{\lambda}_k$. Par suite, nous effectuons une recherche dichotomique sur α_1 qui conduit à la nouvelle valeur $\hat{\alpha}_{1,k+1}$. Ici, nous utilisons $]-1, 1[$ comme intervalle initial de la recherche. Le pseudo-code suivant définit la recherche dichotomique utilisée pour aboutir à $\hat{\alpha}_{1,k+1}$.

```

left <- -1;
right <- 1;
range <- 2;
mid_left <- ( left + previous_alpha ) / 2;
mid_right <- ( right + previous_alpha ) / 2;
while range > 0.001 do
begin
  V_1 <- Phi_n ( alpha_1,k , alpha_2,k , lambda_k );
  V_2 <- Phi_n ( mid_left , alpha_2,k , lambda_k );
  V_3 <- Phi_n ( mid_right , alpha_2,k , lambda_k );

```

```

i <- j such that V_j is min (V_1, V_2, V_3)
case i of
  1 : begin left <- mid_left; right <- mid_right end;
  2 : begin right <- alpha_1,k; alpha_1,k <- mid_left end;
  3 : begin left <- alpha_1,k; alpha_1,k <- mid_right end;
range <- abs ( right - left )
end;

```

Dans la deuxième phase nous fixons $\hat{\alpha}_{1,k+1}$ et $\hat{\lambda}_k$. Par suite, nous effectuons une recherche dichotomique sur α_2 qui nous conduit à la nouvelle valeur $\hat{\alpha}_{2,k+1}$. Ici, nous utilisons aussi $] -1, 1[$ comme intervalle initial de la recherche. Le pseudo-code suivant définit la recherche dichotomique utilisée pour arriver à $\hat{\alpha}_{2,k+1}$.

```

left <- -1;
right <- 1;
range <- 2;
mid_left <- ( left + previous_alpha ) / 2;
mid_right <- ( right + previous_alpha ) / 2;
while range > 0.001 do
begin
  V_1 <- Phi_n ( alpha_1,k+1 , alpha_2,k , lambda_k );
  V_2 <- Phi_n ( alpha_1,k+1 , mid_left , lambda_k );
  V_3 <- Phi_n ( alpha_1,k+1 , mid_right , lambda_k );
  i <- j such that V_j is min (V_1, V_2, V_3)
  case i of
    1 : begin left <- mid_left; right <- mid_right end;
    2 : begin right <- alpha_2,k; alpha_2,k <- mid_left end;
    3 : begin left <- alpha_2,k; alpha_2,k <- mid_right end;
  range <- abs ( right - left )
end;

```

Dans la dernière phase, nous fixons $\hat{\alpha}_{k+1}$ et $\hat{\alpha}_{2,k+1}$. Par suite, nous effectuons une recherche dichotomique sur λ qui nous conduit à la nouvelle valeur $\hat{\lambda}_{k+1}$. Ici nous proposons $\left[\hat{\lambda}_0 - 5 |\hat{\lambda}_0|, \hat{\lambda}_0 + 5 |\hat{\lambda}_0| \right]$ comme intervalle initial de la recherche, lequel nous semble assez large pour couvrir la plupart des situations. Le pseudo-code suivant définit la recherche dichotomique utilisée pour trouver à $\hat{\lambda}_{k+1}$.

```

left <- lambda_0 - 5 |lambda_0|;
right <- lambda_0 + 5 |lambda_0|;

```

```

range <- 10|lambda_0|;
mid_left <- ( left + previous_lambda ) / 2;
mid_right <- ( right + previous_lambda ) / 2;
while range > 0.001 do
  begin
    V_1 <- Phi_n ( alpha_1,k+1 , alpha_2,k+1 , lambda_k );
    V_2 <- Phi_n ( alpha_1,k+1 , alpha_2,k+1 , mid_left );
    V_3 <- Phi_n ( alpha_1,k+1 , alpha_2,k+1 , mid_right );
    i <- j such that V_j is min ( V_1, V_2, V_3 )
    case i of
      1 : begin left <- mid_left; right <- mid_right end;
      2 : begin right <- lambda_k; lambda_k <- mid_left end;
      3 : begin left <- lambda_k; lambda_k <- mid_right end;
    range <- abs ( right - left )
  end;

```

Ainsi, c'est la fin de la $(k+1)$ ^{ème} itération. L'algorithme s'arrête quand les résultats de deux itérations consécutives sont très proches. En d'autres termes, nous arrêtons à la i ^{ème} itération si

$$d(\hat{\theta}_i, \hat{\theta}_{i+1}) = \max \left\{ |\hat{\alpha}_{1,i} - \hat{\alpha}_{1,i+1}|, |\hat{\alpha}_{2,i} - \hat{\alpha}_{2,i+1}|, |\hat{\lambda}_i - \hat{\lambda}_{i+1}| \right\} \leq 0.001. \quad (3.44)$$

En général, ce critère d'arrêt est satisfait après peu d'itérations. À la fin des itérations, nous obtenons le vecteur $(\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\lambda}_n)$ qui minimise la fonction objectif (de contraste) φ_n , définie par (3.10).

3.3.3 Une expérience de simulation

Notre objectif est d'illustrer les performances de la méthode choisie pour calculer $\hat{\theta}_n$. Pour cela, nous simulons 500 observations du processus RINAR(4), avec une vraie valeur $\theta_0 = (\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_4^*, \lambda^*)$ fixée. Pour obtenir un bruit centré à valeurs dans \mathbb{Z} , nous considérons deux suites ξ_t et η_t indépendantes de variables i.i.d. suivant une loi de Poisson de paramètre μ , et nous posons $\varepsilon_t = \xi_t - \eta_t$.

Puis, en utilisant l'algorithme indiqué précédemment, nous calculons $\hat{\theta}_n$.

Après 500 répétitions indépendantes, nous calculons la moyenne et l'écart-type des suites des estimations obtenues $\{\hat{\alpha}_{1n,i}, 1 \leq i \leq 500\}$, $\{\hat{\alpha}_{2n,i}, 1 \leq i \leq 500\}$, $\{\hat{\alpha}_{3n,i}, 1 \leq i \leq 500\}$, $\{\hat{\alpha}_{4n,i}, 1 \leq i \leq 500\}$ et $\{\hat{\lambda}_{n,i}, 1 \leq i \leq 500\}$, et dessinons leurs histogrammes.

Nous prenons $\theta_0 = (\frac{3}{25}, \frac{3}{8}, \frac{1}{5}, -\frac{1}{4}, \frac{5}{2})$. Notons que cette valeur de θ_0 a été utilisée dans la

Section 3.2.2. Ainsi, rappelons qu'ici

$$\nu_0 = \frac{1}{40} = 0.025, \quad I_0 = [2.5, 2.525[\quad \text{et} \quad |I_0| = 0.025.$$

Le tableau suivant nous donne la moyenne et l'écart-type des suites des estimations obtenues.

θ_0	moyenne	écart-type
$\alpha_1^* = 0.12$	0.1204	0.0473
$\alpha_2^* = 0.375$	0.3687	0.0439
$\alpha_3^* = 0.2$	0.1932	0.0425
$\alpha_4^* = -0.25$	-0.2472	0.0454
$\lambda^* = 2.5$	2.538	0.2766

Les figures suivantes nous donnent les histogrammes des estimations obtenues.

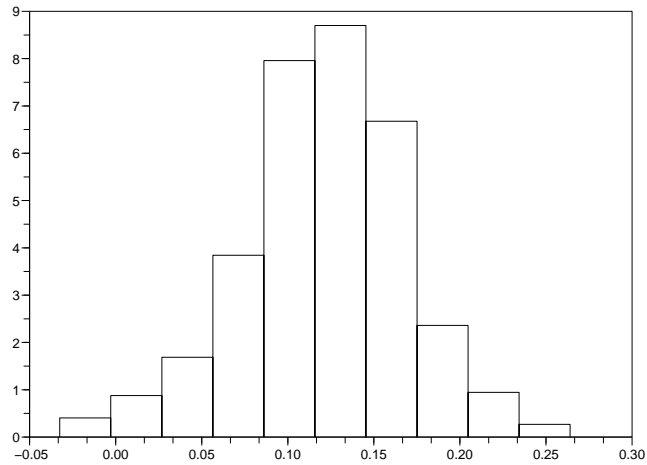


FIG. 3.1 – L'histogramme des $(\hat{\alpha}_{1n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_1^* = 0.12$.

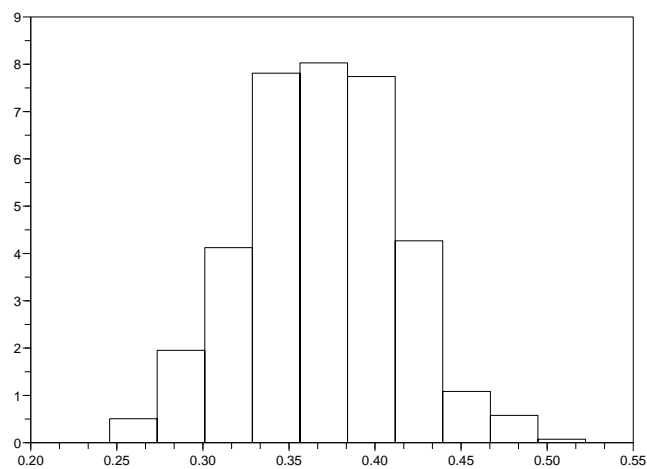


FIG. 3.2 – L’histogramme des $\left(\hat{\lambda}_{2n,i}\right)_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_2^* = 0.375$.

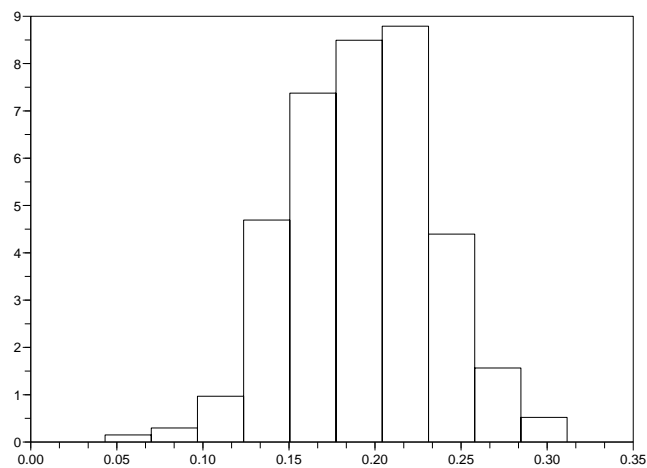


FIG. 3.3 – L’histogramme des $\left(\hat{\lambda}_{3n,i}\right)_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_3^* = 0.2$.

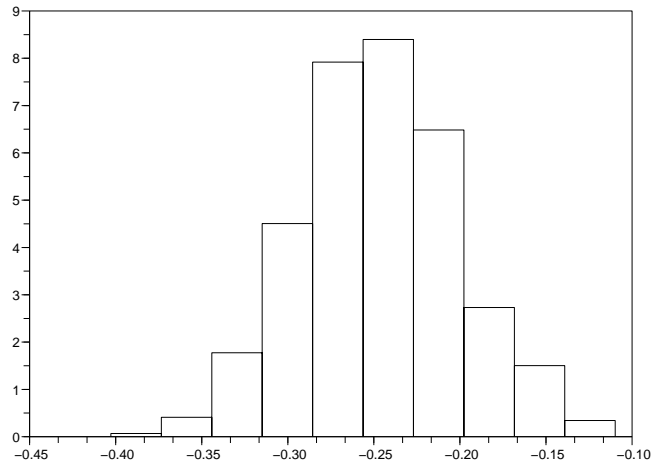


FIG. 3.4 – L’histogramme des $(\hat{\lambda}_{4n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_4^* = -0.25$.

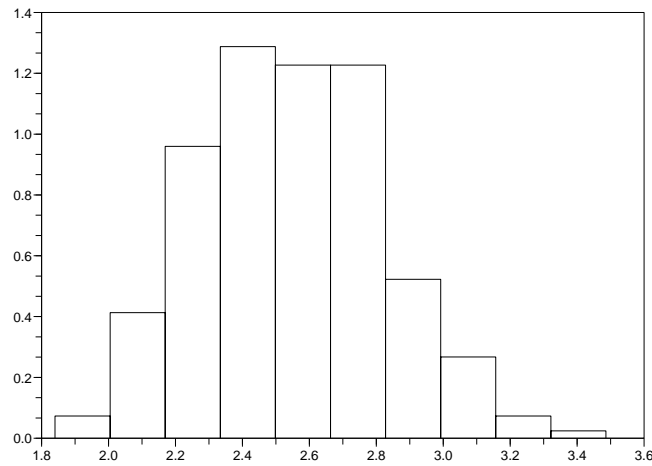


FIG. 3.5 – L’histogramme des $(\hat{\lambda}_{4n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda^* = 2.5$.

Notons que, dans cette étude de simulation, nous avons considéré que le bruit est grand (c.à.d. $\mu \gg \frac{1}{2}$). De plus, nous avons aussi $\lambda^* \gg \frac{1}{2}$. Notons que dans ce cas, l’estimateur de Yule-Walker $\hat{\theta}_0$ des paramètres, qui constitue le point de départ de notre algorithme, est considéré comme significatif.

En d'autres termes, l'estimateur des moindres carrés $\hat{\theta}_n$ calculé suivant notre algorithme doit être très proche de $\hat{\theta}_0$.

3.4 Analyse des données de Fürth

Dans cette section, nous présentons une application à des données réelles. La série de Fürth, initialement publié dans [28], contient 505 comptes qui représentent le nombre des piétons observés toutes les 5 secondes, voir Figure 3.6. Cet ensemble de données est bien connu dans la littérature des processus de branchement, voir Mills & Seneta (1989, [54]). Les comptes varient entre 0 et 7. La moyenne des observations est 1.59 et sa variance est 1.51. La Figure 3.7 montre l'ACF et PACF empiriques des données. Notons que les coefficients d'autocorrélation partiels du premier et second ordre sont significatifs. De plus, le premier est positif tandis que le second est de signe négatif.

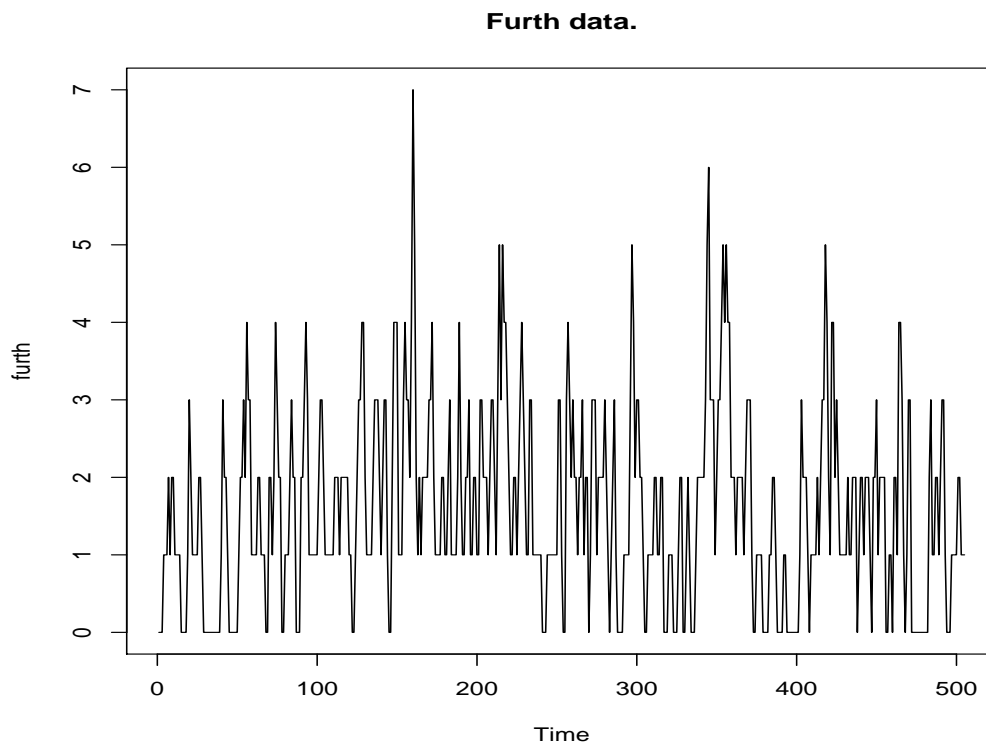


FIG. 3.6 – Les données de Fürth.

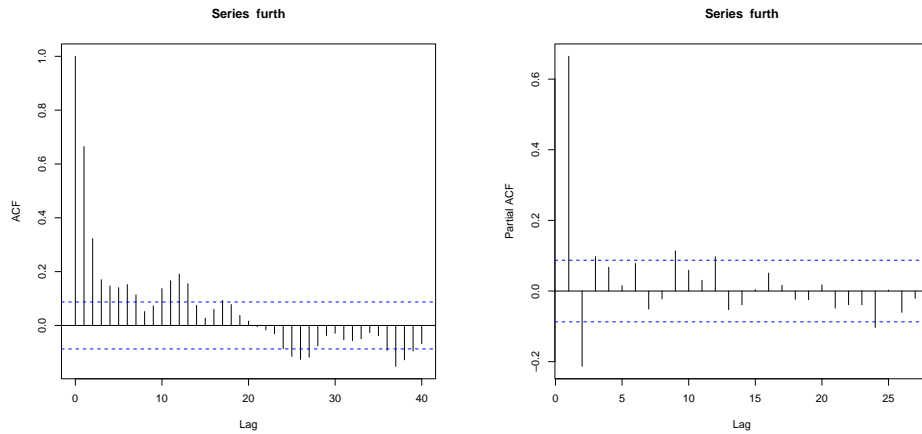


FIG. 3.7 – L'ACF et le PACF de la série Fürth.

3.4.1 Le modèle INAR(2) de Jung & Termayne

Pour analyser cette série, Jung & Termayne (2006, [40]) ont proposé le modèle INAR(2) défini par

$$X_t = a_1 \circ X_{t-1} + a_2 \circ X_{t-2} + b + \varepsilon_t,$$

où $a_1, a_2 \in [0, 1[$ et (ε_t) une suite de variables i.i.d. à valeurs dans \mathbb{N} .

Les auteurs ont considéré les deux spécifications de la classe INAR d'ordre supérieur à un, étudiées dans le Chapitre 1. Rappelons que, ces spécifications dépendent du choix des séries de comptage associées à a_1 et a_2 .

- Pour la spécification de Al-Osh & Alzaid (1990, [6]), notée INAR(2)-AA, ils obtiennent

$$\hat{a}_1 = 0.664, \hat{a}_2 = -0.119, \text{ et } \hat{b} = 0.723.$$

- Pour la spécification de Du & Li (1991, [21]), notée INAR(2)-DL, ils trouvent

$$\hat{a}_1 = 0.808, \hat{a}_2 = -0.214, \text{ et } \hat{b} = 0.646.$$

Aucune des spécifications utilisées ne rapporte un modèle satisfaisant, parce que la valeur estimée de a_2 est négative ce qui n'est pas autorisé dans un modèle INAR.

D'ailleurs, Jung & Termayne ont proposé aussi le modèle INMA(2) introduit par McKenzie (1988, [52]) et défini par

$$X_t = b_1 \circ \varepsilon_{t-1} + b_2 \circ \varepsilon_{t-2} + b + \varepsilon_t, \text{ où } b_1, b_2 \in [0, 1[.$$

Ainsi, les estimations de Yule-Walker des paramètres sont

$$\hat{b}_1 = 1.008, \hat{b}_2 = 0.961, \text{ et } \hat{b} = 0.536.$$

Une difficulté principale ici est que l'estimation du premier paramètre d'amincissement b_1 n'appartient pas à l'espace des paramètres admissibles $[0, 1[$.

3.4.2 Ajustement à un modèle RINAR(2)

En premier lieu, en utilisant le logiciel R, nous allons chercher un modèle autorégressif réel pour analyser les données de Fürth. D'après l'ACF, voir Figure 3.7, nous constatons que les coefficients du premier et second ordre sont plus significatifs que les autres. D'après le PACF, voir Figure 3.7, nous pouvons dire que seulement les coefficients du premier et second ordre sont significatifs.

D'autres part, nous proposons un modèle AR(2)

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + b + \varepsilon_t.$$

Les estimations Yule-Walker des paramètres sont 0.808 pour a_1 (*s.e.*0.0434), -0.214 pour a_2 (*s.e.*0.0435), et 0.646 pour b (les mêmes estimations des paramètres du modèle INAR(2)-DL). Notons que, a_1 et a_2 sont significatifs. De plus, la valeur AIC est égale à 1328.99. Maintenant, nous considérons le modèle AR(3) suivant

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + b + \varepsilon_t.$$

Les estimations Yule-Walker des paramètres sont 0.828 pour a_1 (*s.e.*0.0442), -0.292 pour a_2 (*s.e.*0.0561), 0.0986 pour a_3 (*s.e.*0.0443), et 0.578 pour b . Notons que, a_3 n'est pas significatif. De plus, la valeur AIC associée au modèle AR(3) est 1326.05.

Cette dernière est très proche de la valeur AIC associée au modèle précédent. Ainsi, le meilleur modèle autorégressif réel pour analyser les données présentes est d'ordre 2.

Par suite, pour analyser les données de Fürth, nous proposons le modèle RINAR(2) suivant

$$X_t = \langle \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda \rangle + \varepsilon_t. \quad (3.45)$$

Afin d'examiner les performances de ce modèle, nous réservons les 400 premières observations pour estimer les paramètres et les 105 dernières pour comparer les résultats de prévision.

Pour estimer les paramètres du modèle, nous considérons l'estimateur des moindres carrés

$\hat{\theta}_n = (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\lambda}_n)$ défini par (3.9) et (3.10). Pour calculer $\hat{\theta}_n$, nous utilisons l'algorithme proposé dans la section 3.3.

Ainsi, nous obtenons

$$\hat{\theta}_n = (\hat{\alpha}_{1,n}, \hat{\alpha}_{2,n}, \hat{\lambda}_n) = (0.818, -0.23, 0.697).$$

Soit T un instant fixé. La prévision à un pas \hat{X}_{T+1} , basée sur la fonction de régression, est égale

$$\hat{X}_{T+1} = \langle \hat{\alpha}_{1,n} X_T + \hat{\alpha}_{2,n} X_{T-1} + \hat{\lambda}_n \rangle. \quad (3.46)$$

La figure 3.8 nous montre les erreurs de prévision définies par

$$\hat{\varepsilon}_{T+1} = X_{T+1} - \hat{X}_{T+1}, \quad 400 \leq T \leq 504$$

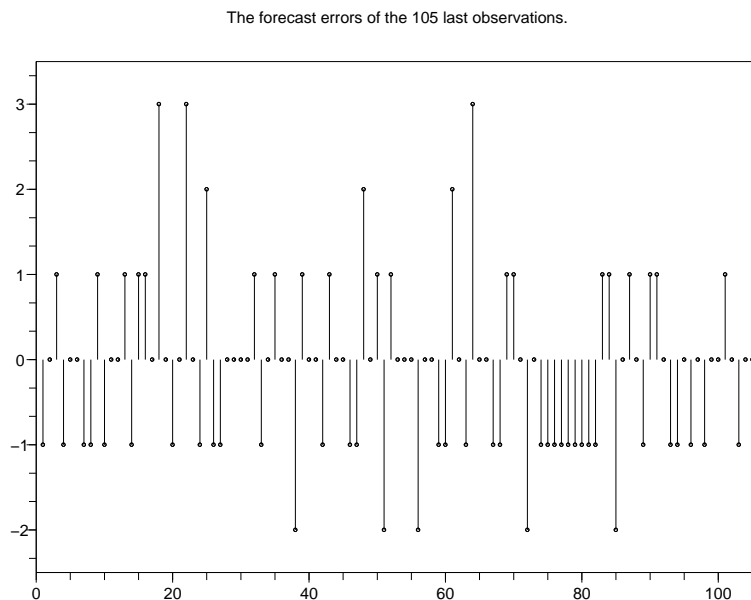


FIG. 3.8 – Les erreurs de prévision $\hat{\varepsilon}_{T+1} = X_{T+1} - \hat{X}_{T+1}$, $400 \leq T \leq 504$, provenant du modèle RINAR(2).

Alors, l'erreur moyenne absolue est

$$EMA = \frac{1}{105} \sum_{T=400}^{504} |\hat{\varepsilon}_{T+1}| = 0.743. \quad (3.47)$$

D'autre part, les résidus du modèle sont définis par

$$\tilde{\varepsilon}_t = X_t - \langle \hat{\alpha}_{1,n} X_{t-1} + \hat{\alpha}_{2,n} X_{t-2} + \hat{\lambda}_n \rangle, \quad 2 \leq t \leq 400. \quad (3.48)$$

La valeur EMA n'est suffisamment satisfaisante, ce qui implique que d'autres modèles de la classe RINAR pourraient être envisagés. Cependant, l'utilisation du modèle RINAR(2) est plus naturelle que les modèles précédemment proposés par Jung & Termayne (2006, [40]). En particulier, le modèle estimé est compatible avec le domaine de définition des paramètres.

Finalement, pour vérifier la validité du modèle proposé, nous simulons 1000 observations du processus RINAR(2) avec $\theta_0 = \hat{\theta}_n$, notamment

$$X_i = \langle 0.818 X_{i-1} - 0.23 X_{i-2} + 0.697 \rangle + \varepsilon_i.$$

Ici, le bruit (ε_i) est défini comme la différence de deux variables aléatoires i.i.d. suivant une loi de Poisson de paramètre μ . Ainsi, nous avons $\mathbb{E}\varepsilon_i = 0$ et $V(\varepsilon_i) = 2\mu$.

Le paramètre μ est déterminé de la manière suivante. La moyenne quadratique des résidus est

$$\frac{1}{309} \sum_{t=2}^{400} \tilde{\varepsilon}_t = 0.834.$$

Ainsi, nous prenons

$$\mu = \frac{0.834}{2} = 0.417.$$

Les figures suivantes nous donnent l'ACF et le PACF empiriques de la série simulée.

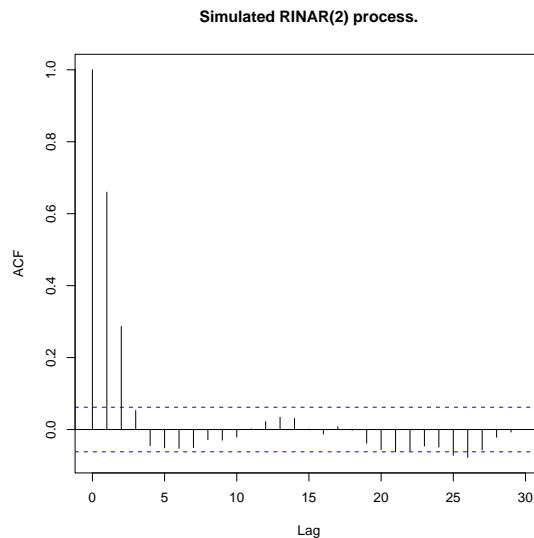


FIG. 3.9 – ACF empirique de la série simulée du modèle RINAR(2), avec $\theta_0 = \hat{\theta}_n = (0.818, -0.23, 0.697)$.

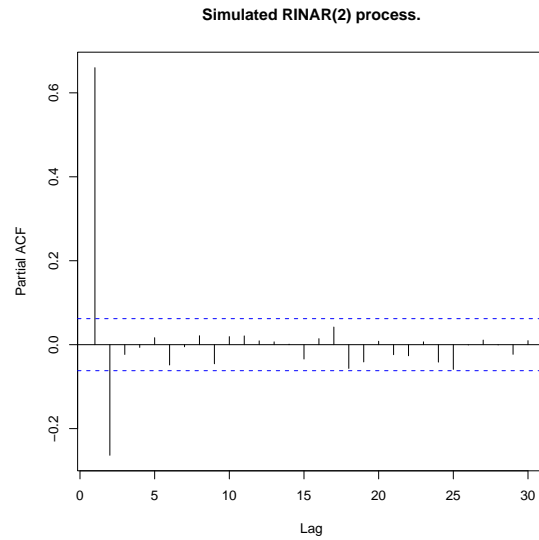


FIG. 3.10 – PACF empirique de la série simulée du modèle RINAR(2), avec $\theta_0 = \hat{\theta}_n = (0.818, -0.23, 0.697)$.

Ainsi, il est clair que les premiers pas de l'ACF (resp. PACF) empiriques de la série simulée du processus RINAR(2) sont très proches de ceux de la série Fürth, voir Figure 3.7. le modèle proposé peut reproduire l'ACF et PACF empirique des données réelles.

Chapitre 4

Le modèle RINVAR(1)

Sommaire

4.1	Stationnarité et ergodicité du modèle RINVAR(1)	118
4.2	Estimation des paramètres	122
4.2.1	La consistance forte de l'estimateur des moindres carrés quand chaque ligne de la matrice M_0 possède un coefficient irrationnel	127
4.2.2	La consistance forte de l'estimateur des moindres carrés quand tous les coefficients de la matrice M_0 sont rationnels	130
4.2.3	La consistance forte de l'estimateur des moindres carrés quand au moins une ligne de la matrice M_0 possède au moins un coefficient irrationnel	134
4.3	Méthode numérique pour calculer $\hat{\theta}_n$	137
4.4	Analyse des taux de variation et les indices de récolte annuels de la population Suédoise	138
4.4.1	L'approche de McCleary & Hay	139
4.4.2	Ajustement à un modèle RINVAR(1)	140
4.4.3	Comparaison et commentaires	142

Dans ce chapitre nous étudions le modèle RINVAR, l'extension directe et naturelle du processus RINAR(p) examiné dans le Chapitre 3, pour analyser les séries chronologiques multivariées. En particulier, nous nous concentrons sur le modèle du premier ordre, noté RINVAR(1) (pour rounded integer-valued vectorial autoregressive). Ce dernier a été étudié par Kachour (2009, [41]).

Dans le premier paragraphe, nous donnons les conditions de stationnarité et d'ergodicité du processus. Dans le deuxième paragraphe, nous proposons l'estimateur des moindres

carrés pour estimer les paramètres du modèle RINVAR. Puis, nous examinons le problème d'identifiabilité, du à l'opérateur d'arrondi, et la consistance de ce dernier. Dans le troisième paragraphe, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Finalement, dans le quatrième paragraphe, nous présentons une application réelle traitée avec un modèle RINVAR(1). Notons que dans ce chapitre nous gardons les mêmes notations et définitions que dans les Chapitres 2 et 3.

4.1 Stationnarité et ergodicité du modèle RINVAR(1)

Définition 4.1.1. Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $d \geq 2$. Le processus $\{X_t = (X_{1,t}, \dots, X_{d,t}), t \in \mathbb{Z}\}$ est dit un processus RINVAR(p) si pour tout t , il possède la présentation suivante

$$X_t = \left\langle \sum_{j=1}^p M_j X_{t-j} + \lambda \right\rangle + \varepsilon_t, \quad (4.1)$$

où $(M_j)_{j=1, \dots, p}$ est une collection finie de matrice d'ordre d (c.à.d. $M_j \in \mathcal{M}_d(\mathbb{R}), \forall j = 1, \dots, p$), $\lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$, $\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près, et (ε_t) est une suite de variables aléatoires vectorielles centrées i.i.d., à valeurs dans \mathbb{Z}^d , définie sur $(\Omega, \mathcal{A}, \mathbb{P})$.

Nous considérons que la série chronologique vectorielle de dimension $d \geq 2$ suivante

$$X_t = \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{d,t} \end{pmatrix}, \quad (4.2)$$

est générée par le processus RINVAR(1) défini par

$$X_t = \langle M X_{t-1} + \lambda \rangle + \varepsilon_t, \quad (4.3)$$

où

$$M = (\alpha_{ij}) \text{ pour } 1 \leq i, j \leq d, \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_d \end{pmatrix}, \text{ et } \varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{d,t} \end{pmatrix}. \quad (4.4)$$

De plus, (ε_t) est une suite de variables aléatoires vectorielles centrées i.i.d., à valeurs dans \mathbb{Z}^d et de matrice de variance-covariance Γ .

Soit $x = (x_1, \dots, x_d)^\tau \in \mathbb{R}^d$. Rappelons que l'arrondi du vecteur x , noté $\langle x \rangle$, est défini par

$$\langle x \rangle = (\langle x_1 \rangle, \dots, \langle x_d \rangle)^\tau \in \mathbb{Z}^d.$$

De plus, nous avons $\|x\|_1 = |x_1| + \dots + |x_d|$ et $(\|x\|_2)^2 = |x_1|^2 + \dots + |x_d|^2$.

Comme le bruit (ε_t) est une suite de variables aléatoires i.i.d., le processus (X_t) défini par (4.3), forme une chaîne de Markov homogène avec un espace d'états $E = \mathbb{Z}^p$ et une probabilité de transition

$$\pi(x, y) = \mathbb{P}(\varepsilon_1 = y - \langle M x + \lambda \rangle) \quad \forall x = (x_j), y = (y_j) \in E. \quad (4.5)$$

Notons que, pour toute mesure μ et une fonction g sur E , nous avons

$$\mu(g) = \int g(x) d\mu(x) = \sum_{i \in E} g(i) \mu(i).$$

De plus, pour tout $n \in \mathbb{N}^*$ et $x \in E$, nous avons

$$\begin{aligned} (\pi^n g)(x) &= \mathbb{E}[g(X_{t+n}) \mid X_t = x], \quad \forall t \in \mathbb{Z} \\ &= \mathbb{E}_x[g(X_n)], \end{aligned}$$

où $\mathbb{E}_x[\cdot]$ représente l'espérance conditionnelle $\mathbb{E}[\cdot \mid X_0 = x]$.

Soit $k \geq 2$ fixé. Nous définissons sur E la fonction ψ par

$$x = (x_1, \dots, x_d)^\tau \mapsto \psi(x) = \begin{pmatrix} |x_1|^k \\ \vdots \\ |x_d|^k \end{pmatrix}. \quad (4.6)$$

Proposition 4.1.1. *Supposons que :*

1. *La chaîne de Markov (X_t) est irréductible ;*
2. *$\mathbb{E}[\psi(\varepsilon_t)] < \infty$, c.à.d. $\mathbb{E}|\varepsilon_{j,t}|^k < +\infty$, $1 \leq j \leq d$, où $k \geq 2$;*
3. *Le rayon spectral de $A = |M| = (|\alpha_{ij}|)$ est strictement inférieur à 1.*

Alors,

1. *(X_t) possède une unique mesure de probabilité invariante, notée μ . De plus, μ possède un moment d'ordre k .*
2. *Pour tout $x \in E$ et $f \in L^1(\mu)$ nous avons*

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \longrightarrow \mu(f), \quad \mathbb{P}_x \text{ p.s.}$$

où \mathbb{P}_x représente la probabilité conditionnelle $\mathbb{P}(\cdot \mid x_0 = x)$.

Preuve. nous définissons sur E les fonctions φ et V par

$$x \mapsto \varphi(x) = \begin{pmatrix} |x_1| \\ \vdots \\ |x_p| \end{pmatrix}, \quad x \mapsto V(x) = \left(\sum_{j=1}^p |x_j| \right)^k = (\|x\|_1)^k.$$

Comme V est positive et $\lim_{|x| \rightarrow \infty} V(x) = \infty$, V est donc une fonction de Lyapunov.

Soit " \leq " une relation d'ordre partiel sur \mathbb{R}^d définie par

$$x \leq y \iff |x_j| \leq |y_j|, \quad \forall j = 1, \dots, p,$$

pour $x = (x_j)$ et $y = (y_j)$. Rappelons que, pour tout $a \in \mathbb{R}$, nous avons

$$|\langle a \rangle| = \langle |a| \rangle \leq |a| + \frac{1}{2}.$$

Ainsi, pour tout $j \in \{1, \dots, d\}$, nous obtenons

$$|\langle \sum_{k=1}^d \alpha_{jk} X_{k,t-j} + \lambda_j \rangle + \varepsilon_{j,t}| \leq \sum_{k=1}^d |\alpha_{jk}| |X_{k,t-j}| + |\varepsilon_{j,t}| + |\lambda_j| + \frac{1}{2}.$$

Par conséquent, nous trouvons

$$\varphi(X_t) \leq A \varphi(X_{t-1}) + \zeta_t,$$

où

$$A = |M| = (|\alpha_{ij}|) \text{ et } \zeta_t = \begin{pmatrix} |\varepsilon_{1,t}| + |\lambda_1| + \frac{1}{2} \\ \vdots \\ |\varepsilon_{d,t}| + |\lambda_d| + \frac{1}{2} \end{pmatrix}.$$

Ainsi pour tout $n \in \mathbb{N}^*$, nous trouvons par itération

$$\varphi(X_{t+n}) \leq A^n \varphi(X_t) + A^{n-1} \zeta_{t+1} + A^{n-2} \zeta_{t+2} + \dots + A \zeta_{t+n-1} + \zeta_{t+n}.$$

Par suite, comme $\|\cdot\|_1$ est fonction croissante sur (\mathbb{R}^2, \leq) , nous obtenons

$$\|\varphi(X_{t+n})\|_1 \leq \|A^n \varphi(X_t)\|_1 + \|A^{n-1}\|_1 \|\zeta_{t+1}\|_1 + \dots + \|A\|_1 \|\zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1,$$

où $\|\cdot\|_1$ est la norme matricielle associée à la norme $\|\cdot\|_1$. Par conséquent,

$$((\pi^n V)(x))^{\frac{1}{k}} \leq \left(\mathbb{E}_x [\|A^n \varphi(X_0)\|_1 + \|A^{n-1}\|_1 \|\zeta_{t+1}\|_1 + \dots + \|A\|_1 \|\zeta_{t+n-1}\|_1 + \|\zeta_{t+n}\|_1]^k \right)^{\frac{1}{k}}$$

Notons que, pour tout $j \in \{1, \dots, n\}$, $\|\zeta_{t+j}\|_1 = \|\varepsilon_{t+j}\|_1 + \|\lambda\|_1 + \frac{d}{2}$.

Ainsi, comme Θ est compact et $\mathbb{E}|\varepsilon_{i,t}|^k < +\infty$, $1 \leq i \leq d$, nous avons

$$b = \left(\mathbb{E} [\|\zeta_{t+j}\|_1]^k \right)^{\frac{1}{k}} < \infty, \quad \forall j \in \{1, \dots, n\}.$$

Par conséquent, d'après l'inégalité de Minkowski, nous obtenons

$$((\pi^n V)(x))^{\frac{1}{k}} \leq \|A^n\|_1 \|x\|_1 + S_n(A) \|x\|_1,$$

où

$$S_n(A) = \|A^{n-1}\|_1 + \dots + \|A\|_1 + 1,$$

Comme $\rho(A) < 1$, il existe $n_0 \in \mathbb{N}^*$ tel que $\forall N \geq n_0$, nous avons $\|A^N\|_1 = \alpha' < 1$. Par suite, d'après le critère de Cauchy, nous obtenons que la série $\sum \|A^i\|_1$ est convergente. Ainsi, nous trouvons

$$((\pi^N V)(x))^{\frac{1}{k}} \leq \alpha' \|x\|_1 + \beta = \|x\|_1 \left(\alpha' + \frac{\beta}{\|x\|_1} \right)$$

et par suite

$$\frac{(\pi^N V)(x)}{V(x)} \leq \left(\alpha' + \frac{\beta}{\|x\|_1} \right)^k,$$

où $\alpha' < 1$ et $0 < \beta < \infty$. Par conséquent, il existe $N \in \mathbb{N}^*$ ($N \geq n_0$) tel que

$$\limsup_{\|x\|_1 \rightarrow \infty} \frac{\pi^N V(x)}{V(x)} \leq |\alpha'|^k < 1.$$

Ceci implique que

$$\lim_{n \rightarrow \infty} \sup \mu_n V \leq l < \infty,$$

où $\mu_n(\cdot)$ est la mesure empirique de la chaîne de Markov (X_t) .

Ainsi, le critère de Lyapunov est satisfait (voir Duflo [22], 1997, Prop. 2.1.6, p. 41).

Alors, (X_t) possède une mesure invariante μ tel que $\mu V < \infty$. De plus, (X_t) est irréductible.

Il en découle, (X_t) est récurrente positive et par suite μ est unique.

La deuxième conclusion de la proposition résulte directement du théorème ergodique classique des chaînes de Markov. □

4.2 Estimation des paramètres

Tout d'abord, nous présentons quelques notations et rappels nécessaires.

Nous associons à la norme $\|\cdot\|_1$ sur \mathbb{R}^d , une norme sur $(\mathbb{R}^d)^2$ définie par

$$|z| = \|y_1\|_1 + \|y_2\|_1, \quad \forall z = (y_1, y_2) \in (\mathbb{R}^d)^2.$$

Soit $\theta = (M, \lambda) \in \Theta \subset \mathcal{M}_d([-1, 1]) \times \mathbb{R}^d$, où $M = (\alpha_{ij})_{1 \leq i, j \leq d}$ et $\lambda = (\lambda_1, \dots, \lambda_d)^\tau \in \mathbb{R}^d$.

Notons que l'espace des paramètres Θ est compact.

Soit $x = (x_1, \dots, x_d)^\tau \in \mathbb{R}^d$. Pour tout $i \in \{1, \dots, d\}$, nous notons

$$\alpha_i = (\alpha_{i1}, \dots, \alpha_{id}) \text{ et } f_i(x; \alpha_i, \lambda_i) = \left\langle \sum_{j=1}^d \alpha_{ij} x_j + \lambda_i \right\rangle. \quad (4.7)$$

Par suite, nous avons

$$M = (\alpha_1, \dots, \alpha_d)^\tau. \quad (4.8)$$

Rappelons que,

$$\langle x \rangle = (\langle x_1 \rangle, \dots, \langle x_d \rangle)^\tau \in E = \mathbb{Z}^d.$$

Nous définissons la fonction de régression du modèle RINVAR(1) par

$$F(x; \theta) = \langle Mx + \lambda \rangle, \quad x \in E. \quad (4.9)$$

Ainsi, RINVAR(1) peut être écrit sous la forme suivante

$$X_t = F(X_{t-1}; \theta) + \varepsilon_t = \begin{pmatrix} f_1(X_{t-1}; \alpha_1, \lambda_1) \\ \vdots \\ f_d(X_{t-1}; \alpha_d, \lambda_d) \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{d,t} \end{pmatrix}. \quad (4.10)$$

Donc, Pour tout $i \in \{1, \dots, d\}$, nous avons

$$X_{i,t} = f_i(X_{t-1}; \alpha_i, \lambda_i) = \left\langle \sum_{j=1}^d \alpha_{ij} X_{j,t-1} + \lambda_i \right\rangle + \varepsilon_{i,t}. \quad (4.11)$$

Soit $\theta_0 = (M_0, \lambda_0)$ la vraie valeur du paramètre θ du modèle RINVAR(1),

avec $\lambda_0 = (\lambda_1^*, \dots, \lambda_d^*)^\tau$ et $M_0 = (\alpha_{ij}^*)_{1 \leq i, j \leq d} = (\alpha_1^*, \dots, \alpha_d^*)^\tau$, où $\alpha_i^* = (\alpha_{i1}^*, \dots, \alpha_{id}^*)$

pour tout $i \in \{1, \dots, d\}$. Soit \mathbb{P}_{θ_0} la distribution de la probabilité de la chaîne (Y_t) sous

le vrai modèle. De plus, toute convergence $\xrightarrow{p.s.}$ signifie une convergence *p.s.* sous $\mathbb{P}_{\theta_0, x}$, ce

qui se tient indépendamment de l'état initial x .

Soient X_0, X_1, \dots, X_n des observations du processus RINVAR(1).

Pour l'estimation du paramètre θ , nous considérons l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \varphi_n(\theta), \quad (4.12)$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n (\|X_t - F(X_{t-1}; \theta)\|_2)^2. \quad (4.13)$$

Soit l'hypothèse suivante.

Hypothèse [H3]

1. Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible.
2. $\mathbb{E}[\psi(\varepsilon_t)] < \infty$, c.à.d. $\mathbb{E}|\varepsilon_{j,t}|^k < +\infty$, $1 \leq j \leq d$, où $k \geq 2$.
3. Le rayon spectral de $A = |M| = (|\alpha_{ij}|)$ est strictement inférieur à 1.
4. L'espace des paramètres Θ , est un compact sous-ensemble de $\mathcal{M}_d([-1, 1]) \times \mathbb{R}^d$.

Supposons que [H3] soit vérifiée. Donc, sous \mathbb{P}_{θ_0} et d'après la Proposition 4.1.1, (X_t) possède une unique mesure invariante μ_{θ_0} telle que $\mu_{\theta_0}(|\cdot|^k) < \infty$ avec $k \geq 2$.

Il en résulte que la chaîne double (Y_t) avec $Y_t = (X_{t-1}, X_t)$ possède des propriétés similaires.

En particulier, la chaîne (Y_t) possède aussi une unique mesure invariante σ_{θ_0} tel que $\sigma_{\theta_0}(|\cdot|^k) < \infty$, où $|\cdot|$ est la norme sur $(\mathbb{R}^d)^2$ associée à la norme $\|\cdot\|_1$ (définie sur \mathbb{R}^d).

Soient les fonctions :

$$g_\theta(z) = (\|y - F(x; \theta)\|_2)^2, \quad z = (x, y) \in E^2, \quad \theta \in \Theta, \quad (4.14)$$

et

$$K(\theta) = \sigma_{\theta_0} g_\theta, \quad \theta \in \Theta. \quad (4.15)$$

Soit \mathbb{P}_n la mesure empirique générée par les observations Y_1, \dots, Y_n

$$\mathbb{P}_n(y) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{Y_i=y}, \quad y = (x, z) \in E^2 = \mathbb{Z}^2.$$

Par conséquent, d'après (4.13), la fonction de contraste φ_n est égale

$$\varphi_n(\theta) = \mathbb{P}_n g_\theta \quad (4.16)$$

et par suite

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \mathbb{P}_n g_\theta. \quad (4.17)$$

Lemme 4.2.1. *Pour tout $\theta \in \Theta$, nous avons*

- $(\|F(\cdot; \theta)\|_2)^2 \in L^1(\mu_{\theta_0})$.
- $g_\theta \in L^1(\sigma_{\theta_0})$.

Preuve. Dans la preuve suivante, nous notons par c une constante générique dont la valeur exacte peut changer pendant le développement mathématique.

Pour tout $x = (x_1, \dots, x_d)^\tau \in E$, d'après (4.10), nous avons

$$\|F(x; \theta)\|_1 = \sum_{i=1}^d |f_i(x; \alpha_i, \lambda_i)|,$$

Rappelons que, pour tout $a \in \mathbb{R}$, nous avons $|\langle a \rangle| = \langle |a| \rangle \leq |a| + \frac{1}{2}$.

Ainsi, pour tout $i \in \{1, \dots, d\}$, nous trouvons

$$|f_i(x; \alpha_i, \lambda_i)| = |\langle \sum_{j=1}^d \alpha_{ij} x_j + \lambda_i \rangle| \leq \sum_{j=1}^d |\alpha_{ij}| |x_j| + |\lambda_i| + \frac{1}{2}.$$

Comme Θ est compact, nous avons

$$\|F(x; \theta)\|_1 \leq c (1 + \|x\|_1).$$

Alors, comme sur \mathbb{R}^d les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes, nous obtenons

$$(\|F(x; \theta)\|_2)^2 \leq c (1 + \|x\|_1^2)$$

Donc, comme $\mu_{\theta_0}(\|\cdot\|_1^k) < \infty$ où $k \geq 2$, la première conclusion est vérifiée.

Par suite, pour tout $z = (x, y) \in E^2$, nous avons

$$\begin{aligned} \|y - F(x; \theta)\|_1 &\leq \|y\|_1 + \|F(x; \theta)\|_1 \\ &\leq c (1 + \|x\|_1 + \|y\|_1) \\ &= c (1 + |z|), \end{aligned}$$

où $|\cdot|$ est la norme sur $(\mathbb{R}^d)^2$ associée à la norme $\|\cdot\|_1$ définie sur \mathbb{R}^d .

Il en découle,

$$g_\theta(z) \leq c (1 + |z|^2). \tag{4.18}$$

Donc, comme $\sigma_{\theta_0}(|\cdot|^k) < \infty$ où $k \geq 2$, la deuxième conclusion est vérifiée. \square

Proposition 4.2.1. *Supposons que $[H3]$ soit vérifiée. Donc, pour tout $\theta \in \Theta$, nous avons*

1. $\varphi_n(\theta) \xrightarrow{p.s.} K(\theta);$
2. $K(\theta) - K(\theta_0) = \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2.$

De plus,

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0.$$

Preuve. Rappelons que $\varphi_n(\theta) = \mathbb{P}_n g_\theta$ et $K(\theta) = \sigma_{\theta_0} g_\theta$. Donc, comme $\mathbb{P}_n \xrightarrow{p.s.} \sigma_{\theta_0}$ et d'après le Lemme 4.2.1 $g_\theta \in L^1(\sigma_{\theta_0})$, la première conclusion est vérifiée.

Par suite, nous avons

$$\varphi_n(\theta) - \varphi_n(\theta_0) = \frac{A_n}{n} + \frac{B_n}{n},$$

avec

$$A_n = \sum_{t=1}^n (\|\Delta F_{t-1}\|_2)^2, \quad B_n = 2 \sum_{t=1}^n \langle \varepsilon_t \mid \Delta F_{t-1} \rangle,$$

où $\Delta F_{t-1} = F(X_{t-1}; \theta) - F(X_{t-1}; \theta_0)$ et $\langle \cdot \mid \cdot \rangle$ est le produit scalaire associé à la norme $\|\cdot\|_2$. D'après le Lemme 4.2.1, nous avons $(\|F(\cdot; \theta)\|_2)^2 \in L^1(\mu_{\theta_0})$. Ainsi, en utilisant le théorème ergodique sur la chaîne de Markov (X_t) , nous trouvons

$$\frac{A_n}{n} \xrightarrow{p.s.} \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2.$$

Nous posons,

$$M_n = \frac{B_n}{2}.$$

Il est simple de vérifier que (M_n) est une martingale de carré intégrable par rapport à $\mathcal{F} = (\mathcal{F}_n)_{n \geq 0}$ la filtration naturelle associée au processus RINVAR(p), où $\mathcal{F}_n = \sigma(\varepsilon_t, 0 \leq t \leq n)$ pour tout $n \in \mathbb{N}^*$, et \mathcal{F}_0 est la σ -algèbre dégénérée. Le processus croissant associé, noté $([M]_n)$, est défini par

$$[M]_n = \sum_{t=1}^n (\Delta F_{t-1})^\tau \Gamma \Delta F_{t-1},$$

où Γ est la matrice de variance-covariance de ε_t . Ainsi, nous avons

$$\frac{[M]_n}{n} \xrightarrow{p.s.} \mu_{\theta_0} ((\Delta F_{t-1})^\tau \Gamma \Delta F_{t-1}) \geq 0,$$

Il en résulte, $\frac{M_n}{n} \xrightarrow{p.s.} 0$ et par conséquent $\frac{B_n}{n} \xrightarrow{p.s.} 0$ (voir Duflo [22], Théorème 1.3.15, p. 20). Alors, nous avons

$$K(\theta) - K(\theta_0) = \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2.$$

D'autre part, d'après la preuve du Lemme 4.2.1, la classe des fonctions de contraste

$\mathcal{G} = \{g_\theta, \theta \in \Theta\}$ possède une fonction enveloppe, notée A et définie par

$$A(z) = c(1 + |z|^2), \forall z \in E^2$$

où c est une constante et $|\cdot|$ est la norme sur E^2 associée à la norme $\|\cdot\|_2$.

Soit $q > 0$ fixé. Nous posons,

$$p_z = \sigma_{\theta_0}(z) \text{ et } p_z^n = \mathbb{P}_n(z).$$

Alors, nous obtenons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \leq \sum_{|z| < q} A(z) |p_z^n - p_z| + (\mathbb{P}_n + \sigma_{\theta_0})(A(z) \mathbb{1}_{|z| > q}).$$

Comme $p_z^n \rightarrow p_z$ presque sûrement, nous trouvons

$$\limsup_{n \rightarrow \infty} \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g| \leq 0 + 2\sigma_{\theta_0}(A(z) \mathbb{1}_{|z| > q}), \text{ p.s.}$$

En prenant $q \nearrow \infty$, nous obtenons presque sûrement,

$$\limsup_n \sup_{\theta \in \Theta} |(\mathbb{P}_n - \sigma_{\theta_0})g| = 0.$$

□

Comme pour le modèle RINAR(p), l'identifiabilité du processus RINVAR(1) n'est pas standard à cause de l'opérateur d'arrondi. Dans les paragraphes suivants nous traitons ce problème. Ainsi, nous distinguons trois cas dépendants des coefficients α_{ij}^* de la matrice M_0 .

Nous définissons sur l'espace des paramètres Θ la distance d par

$$d(\theta, \theta') = \max \left\{ |\alpha_{ij} - \alpha'_{ij}|, 1 \leq i, j \leq d; |\lambda_k - \lambda'_k|, 1 \leq k \leq d \right\}, \quad (4.19)$$

où $\theta = (M, \lambda)$ et $\theta' = (M', \lambda')$ avec $M = (\alpha_{ij})$, $\lambda = (\lambda_j)$, $M' = (\alpha'_{ij})$ et $\lambda' = (\lambda'_j)$.

Quelques rappels sont nécessaires. Soient $x = (x_1, \dots, x_d)^\tau \in E$ et $\theta = (M, \lambda) \in \Theta$, avec $M = (\alpha_{ij})_{1 \leq i, j \leq d}$ et $\lambda = (\lambda_1, \dots, \lambda_d)^\tau$. Pour tout $i \in \{1, \dots, d\}$, soient λ_i la $i^{\text{ème}}$ composante du vecteur λ et $\alpha_i = (\alpha_{i1}, \dots, \alpha_{id})$ la $i^{\text{ème}}$ ligne de la matrice M .

De plus,

$$f_i(x; \alpha_i, \lambda_i) = \left\langle \sum_{j=1}^d \alpha_{ij} x_j + \lambda_i \right\rangle \text{ et par suite } F(x; \theta) = \begin{pmatrix} f_1(x; \alpha_1, \lambda_1) \\ \vdots \\ f_d(x; \alpha_d, \lambda_d) \end{pmatrix}.$$

4.2.1 La consistance forte de l'estimateur des moindres carrés quand chaque ligne de la matrice M_0 possède un coefficient irrationnel

Dans ce paragraphe, nous supposons que pour tout $i \in \{1, \dots, d\}$, il existe au moins $j \in \{1, \dots, d\}$ tel que $\alpha_{ij} \in \mathbb{R} \setminus \mathbb{Q}$.

Proposition 4.2.2. *Soit $i \in \{1, \dots, d\}$ fixé. Supposons que [H3] soit vérifiée. S'il existe au moins $j_0 \in \{1, \dots, d\}$ tel que $\alpha_{ij_0}^*$ est irrationnel. Alors,*

$$f_i(x; \alpha_i, \lambda_i) = f_i(x; \alpha_i^*, \lambda_i^*), \quad \forall x \in E \iff (\alpha_i, \lambda_i) = (\alpha_i^*, \lambda_i^*).$$

Preuve. Nous nous référons à la preuve de la Proposition 3.2.2. □

Corollaire 4.2.1. *Supposons que [H3] soit vérifiée. Si chaque ligne de la matrice M_0 possède au moins un coefficient irrationnel. Alors, pour tout $\theta \in \Theta$, nous avons*

$$F(x; \theta) = F(x; \theta_0) \quad \forall x \in E \iff \theta = \theta_0.$$

Preuve. Nous avons,

$$F(x; \theta) = F(x; \theta_0), \quad \forall x \in E \iff f_i(x; \alpha_i, \lambda_i) = f_i(x; \alpha_i^*, \lambda_i^*), \quad \forall i \in \{1, \dots, d\}, x \in E.$$

Ici, pour tout $i \in \{1, \dots, d\}$, il existe au moins $j_i \in \{1, \dots, d\}$ tel que α_{ij_i} est irrationnel. Ainsi, d'après la Proposition 4.2.2, nous obtenons

$$\begin{aligned} F(x; \theta) = F(x; \theta_0), \quad \forall x \in E &\iff (\alpha_1, \lambda_1) = (\alpha_1^*, \lambda_1^*), \dots, (\alpha_d, \lambda_d) = (\alpha_d^*, \lambda_d^*) \\ &\iff \theta = \theta_0. \end{aligned}$$

□

Quelques rappels sont nécessaires. Supposons que [H3] soit vérifiée, alors d'après la Proposition 4.2.1, nous avons

$$K(\theta) - K(\theta_0) = \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2 \tag{4.20}$$

La distance d sur l'espace des paramètres Θ est définie par

$$d(\theta, \theta') = \max \left\{ |\alpha_{ij} - \alpha'_{ij}|, 1 \leq i, j \leq d; |\lambda_k - \lambda'_k|, 1 \leq k \leq d \right\}, \quad \forall \theta, \theta' \in \Theta.$$

D'autre part, soit x_0 fixé. Rappelons que $\{\cdot\}$ représente la fonction partie fractionnaire. Ainsi, la fonction $\alpha \rightarrow \{\alpha x_0\}$ définie sur \mathbb{R} et à valeurs dans $[0, 1[$, est continue en α_0 si $\alpha_0 \in \mathbb{R} \setminus \mathbb{Q}$.

Proposition 4.2.3. *Supposons que [H3] soit vérifiée. Si chaque ligne de la matrice M_0 possède au moins un coefficient irrationnel. Alors, pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_\varepsilon} |K(\theta) - K(\theta_0)| > 0,$$

où

$$\Theta_\varepsilon = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}.$$

Preuve. Nous distinguons trois situations pour l'événement Θ_ε . Ainsi, notons que

$$\Theta_\varepsilon = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3, \text{ où}$$

$$\Gamma_1 = \{\theta : \exists i_0, j_0 \in \{1, \dots, d\}, |\alpha_{i_0 j_0} - \alpha_{i_0 j_0}^*| \geq \varepsilon\}, \quad (4.21)$$

$$\Gamma_2 = \{\theta : \forall i, j \in \{1, \dots, d\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, \exists l_0 \in \{1, \dots, d\} |\lambda_{l_0} - \lambda_{l_0}^*| \geq \varepsilon, \langle \lambda_{l_0} \rangle \neq \langle \lambda_{l_0}^* \rangle\}, \quad (4.22)$$

et

$$\Gamma_3 = \{\theta : \forall i, j \in \{1, \dots, d\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, \exists l_0 \in \{1, \dots, d\} |\lambda_{l_0} - \lambda_{l_0}^*| \geq \varepsilon, \langle \lambda_{l_0} \rangle = \langle \lambda_{l_0}^* \rangle\}. \quad (4.23)$$

Nous allons montrer

$$\inf_{\theta \in \Gamma_i} |K(\theta) - K(\theta_0)| > 0, \quad i = 1, 2, 3.$$

L'idée de la preuve est basée sur l'équation (4.20). Alors, le but est de trouver $y_0 \in E$

tel que

$$\|F(y_0; \theta) - F(y_0; \theta_0)\|_2 > 0, \text{ uniformément sur } \Gamma_i, i = 1, 2, 3.$$

– Nous considérons le cas où $\theta \in \Gamma_1$. Pour simplifier les notations, nous supposons que $i_0 = j_0 = 1$. Ainsi, nous avons $|\alpha_{11} - \alpha_{11}^*| \geq \varepsilon$.

Comme Θ est compact, il existe $B > 0$ tel que $|\lambda_i| \leq B, \forall i \in \{1, \dots, d\}$. De plus, il existe $x_0 \in \mathbb{Z}^*$ tel que

$$|x_0| \varepsilon - 2B - 1 > 0.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Ainsi, nous avons

$$\begin{aligned} \|F(y_0; \theta) - F(y_0; \theta_0)\|_1 &= \sum_{i=1}^d |f_i(y_0; \alpha_i, \lambda_i) - f_i(y_0; \alpha_i^*, \lambda_i^*)| \\ &\geq |f_1(y_0; \alpha_1, \lambda_1) - f_1(y_0; \alpha_1^*, \lambda_1^*)| \\ &= |\langle \alpha_{11} \ x_0 + \lambda_1 \rangle - \langle \alpha_{11}^* \ x_0 + \lambda_1^* \rangle| \\ &\geq |\alpha_{11} - \alpha_{11}^*| |x_0| - 2B - 1 \\ &\geq |x_0| \varepsilon - 2B - 1 > 0. \end{aligned}$$

Comme sur \mathbb{R}^d les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes, alors nous avons

$$\|F(y_0; \theta) - F(y_0; \theta_0)\|_2 > 0.$$

– Nous considérons le deuxième cas, $\theta \in \Gamma_2$. Soit $y_0 = 0_E = (0, \dots, 0)^\tau \in E$.

Ainsi, nous avons

$$\begin{aligned} \|F(y_0; \theta) - F(y_0; \theta_0)\|_1 &= \sum_{i=1}^d |\langle \lambda_i \rangle - \langle \lambda_i^* \rangle|, \\ &\geq |\langle \lambda_{l_0} \rangle - \langle \lambda_{l_0}^* \rangle|, \\ &\geq 1. \end{aligned}$$

Comme sur \mathbb{R}^d les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes, alors nous trouvons

$$\inf_{\theta \in \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

– Maintenant, nous traitons le dernier cas, $\theta \in \Gamma_3$. Pour simplifier les notations, nous supposons que $l_0 = 1$. Ainsi, nous avons $|\alpha_{ij} - \alpha_{ij}^*| < \varepsilon$ pour tout $i, j \in \{1, \dots, d\}$, $|\lambda_1 - \lambda_1^*| \geq \varepsilon$ et $\langle \lambda_1 \rangle = \langle \lambda_1^* \rangle$. Par hypothèse, chaque ligne de la matrice M_0 possède au moins un coefficient irrationnel. Supposons que α_{11}^* soit le coefficient irrationnel de la première ligne de M_0 . Ainsi, nous avons

$$|\alpha_{11} - \alpha_{11}^*| < \varepsilon, \quad |\lambda_1 - \lambda_1^*| \geq \varepsilon \text{ et } \langle \lambda_1 \rangle = \langle \lambda_1^* \rangle,$$

avec $\alpha_{11}^* \in \mathbb{R} \setminus \mathbb{Q}$. D'après la preuve de la Proposition 2.4.3, il existe $x_0 \in \mathbb{Z}$ tel que

$$|\langle \alpha_{11} x_0 + \lambda_1 \rangle - \langle \alpha_{11}^* x_0 + \lambda_1^* \rangle| > 0.$$

Soit $y_0 = (x_0, 0, \dots, 0) \in E$. Ainsi, nous obtenons

$$\begin{aligned} \|F(y_0; \theta) - F(y_0; \theta_0)\|_1 &= \sum_{i=1}^d |f_i(y_0; \alpha_i, \lambda_i) - f_i(y_0; \alpha_i^*, \lambda_i^*)| \\ &\geq |f_1(y_0; \alpha_1, \lambda_1) - f_1(y_0; \alpha_1^*, \lambda_1^*)| \\ &= |\langle \alpha_{11} x_0 + \lambda_1 \rangle - \langle \alpha_{11}^* x_0 + \lambda_1^* \rangle| > 0. \end{aligned}$$

□

Théorème 4.2.1. *Supposons que [H3] soit vérifiée. Si chaque ligne de la matrice M_0 possède au moins un coefficient irrationnel. Alors, l'estimateur des moindres carrés $\hat{\theta}_n$ de θ_0 est fortement consistant, c.à.d.*

$$\hat{\theta}_n \xrightarrow{p.s.} \theta_0.$$

Preuve. Rappelons que, d'après la Proposition 4.2.1 et 4.2.3, nous avons

$$\sup_{\theta \in \Theta} |\varphi_n(\theta) - K(\theta)| \xrightarrow{p.s.} 0$$

et

$$\inf_{\theta \in \Theta_\varepsilon} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon = \{\theta : d(\theta, \theta_0) \geq \varepsilon\}$. La conclusion $\hat{\theta}_n \rightarrow \theta_0$ presque sûrement résulte par des arguments standards de la théorie des M-estimateurs (voir par exemple Van Der Vaart (1998, [61]), Théorème 5.7). \square

4.2.2 La consistance forte de l'estimateur des moindres carrés quand tous les coefficients de la matrice M_0 sont rationnels

Dans ce paragraphe, nous supposons que pour tout $i, j \in \{1, \dots, d\}$, nous avons

$$\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}, \quad \text{où } a_{ij} \in \mathbb{Z}, \quad b_{ij} \in \mathbb{N}^* \quad \text{et} \quad a_{ij} \wedge b_{ij} = 1.$$

Pour mieux comprendre ce cas, nous fixons $i \in \{1, \dots, d\}$. Puis, nous supposons que tous les coefficients de la $i^{\text{ème}}$ ligne sont rationnels. Ainsi, pour tout $j \in \{1, \dots, d\}$, nous avons

$$\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}, \quad \text{où } a_{ij} \in \mathbb{Z}, \quad b_{ij} \in \mathbb{N}^* \quad \text{et} \quad a_{ij} \wedge b_{ij} = 1.$$

Quelques rappels du Chapitre 3 sont nécessaires. Soit $y = (x_1, \dots, x_d)^\tau \in E = \mathbb{Z}^d$. Ainsi, nous obtenons

$$\sum_{j=1}^d \alpha_{ij}^* x_j = \frac{1}{\prod_{j=1}^d b_{ij}} \left(\sum_{l=1}^d A_{il} x_l \right), \quad (4.24)$$

où

$$A_{il} = a_{il} \prod_{j=1, j \neq l}^d b_{ij}. \quad (4.25)$$

D'après le théorème de Bézout, nous obtenons

$$A_{i1}\mathbb{Z} + \dots + A_{ip}\mathbb{Z} = d_i\mathbb{Z}, \quad (4.26)$$

où $d_i = A_{i1} \wedge \dots \wedge A_{ip}$ est le P.G.C.D. de A_{i1}, \dots, A_{ip} . Par conséquent, il existe $x \in \mathbb{Z}$ tel que

$$\sum_{i=1}^d \alpha_{ij}^* x_j = \nu_i x, \quad (4.27)$$

avec

$$\nu_i = \frac{d_i}{\prod_{j=1}^d b_{ij}}. \quad (4.28)$$

Notons que le numérateur et le dénominateur de ν_i ne sont pas nécessairement premiers entre eux. Ainsi, nous réécrivons ν_i avec sa forme de fraction irréductible

$$\nu_i = \frac{a_i}{b_i}, \quad (4.29)$$

où $a_i \in \mathbb{Z}$, $b_i \in \mathbb{N}^*$ et $a_{ij} \wedge b_{ij} = 1$. Soit $\theta = (M, \lambda) \in \Theta$, où $M = (\alpha_{ij})$ et $\lambda = (\lambda_1, \dots, \lambda_d)^\tau$. Rappelons que, pour tout $i \in \{1, \dots, d\}$, $\alpha_i = (\alpha_{i1}, \dots, \alpha_{id})$ représente la $i^{\text{ème}}$ ligne de la matrice M et λ_i est la $i^{\text{ème}}$ composante du vecteur λ .

Proposition 4.2.4. *Soit $i \in \{1, \dots, d\}$ fixé. Supposons que $[H3]$ soit vérifiée et pour tout $j \in \{1, \dots, d\}$, $\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}$ où $a_{ij} \in \mathbb{Z}$, $b_{ij} \in \mathbb{N}^*$ et $a_{ij} \wedge b_{ij} = 1$. Alors,*

1. $f_i(y; \alpha_i^*, \lambda_i) = f_i(y; \alpha_i^*, \lambda_i^*), \forall y \in E \iff \langle \nu_i x + \lambda_i \rangle = \langle \nu_i x + \lambda_i^* \rangle, \forall x \in \mathbb{Z}$.
2. $f_i(y; \alpha_i, \lambda_i) = f_i(y; \alpha_i^*, \lambda_i^*), \forall y \in E \iff \alpha_i = \alpha_i^* \text{ et } \lambda_i \in I_{i,0}$,
où $I_{i,0} = \{\lambda \in \mathbb{R}, \langle \lambda \rangle = \langle \lambda_i^* \rangle\}$. De plus, ce dernier est de taille $\frac{1}{b_i}$, où b_i est le dénominateur de ν_i défini par (4.28) et (4.29).

Preuve. Nous nous référons à la preuve de la Proposition 3.2.4 et celle de la Proposition 3.2.5. □

Notons que la localisation de $I_{i,0}$ dépend de la parité de b_i et de la position de la partie fractionnaire de λ_i^* sur l'intervalle $[0, 1[$. Ainsi, nous distinguons quatre cas :

1. Supposons que b_i est pair. Alors, $I_{i,0}$ est défini par (2.65) et (2.66).
2. Supposons que b_i est impair et $\{\lambda_i^*\} \in \left[0, \frac{1}{2b_i}\right[$. Alors, $I_{i,0}$ est défini par (2.67).
3. Supposons que b_i est impair et $\{\lambda_i^*\} \in \left[\frac{1}{2b_i}, \frac{2b_i - 1}{2b_i}\right[$. Alors, $I_{i,0}$ est défini par (2.68) et (2.69).
4. Supposons que b_i est impair et $\{\lambda_i^*\} \in \left[\frac{2b_i - 1}{2b_i}, 1\right[$. Alors, $I_{i,0}$ est défini par (2.70).

Notons que, pour rester compatible avec les notations du chapitre présent, il suffit de remplacer q par b_i et λ_0 par λ_i^* dans les équations précédentes.

Corollaire 4.2.2. *Supposons que $[H3]$ soit vérifiée. Si tous les coefficients de la matrice M_0 sont rationnels. Alors, pour tout $\theta \in \Theta$, nous avons*

$$F(x; \theta) = F(x; \theta_0) \forall x \in E \iff M = M_0 \text{ et } \lambda \in I_0 = (I_{1,0}, \dots, I_{d,0})^\tau.$$

Preuve. Nous avons,

$$F(x; \theta) = F(x; \theta_0), \forall x \in E \Leftrightarrow \text{pour tout } i \in \{1, \dots, d\}, f_i(x; \alpha_i, \lambda_i) = f_i(x; \alpha_i^*, \lambda_i^*), \forall x \in E.$$

Ici, pour tout $i, j \in \{1, \dots, d\}$, nous avons $\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}$ où $a_{ij} \in \mathbb{Z}$, $b_{ij} \in \mathbb{N}^*$ et $a_{ij} \wedge b_{ij} = 1$. Ainsi, d'après la Proposition 4.2.4, nous obtenons

$$\begin{aligned} F(x; \theta) = F(x; \theta_0), \quad \forall x \in E &\Leftrightarrow \alpha_1 = \alpha_1^* \text{ et } \lambda_1 \in I_{1,0}, \dots, \alpha_d = \alpha_d^* \text{ et } \lambda_d \in I_{d,0}, \\ &\Leftrightarrow M = M_0 \text{ et } \lambda \in I_0. \end{aligned}$$

□

Ainsi, I_0 est un vecteur d'intervalles. La localisation de ce dernier dépend de la parité de chaque b_i et de la position de chaque $\{\lambda_i^*\}$.

Ici, nous détaillons le cas où pour tout $i \in \{1, \dots, d\}$, b_i est pair. Donc, $\forall i \in \{1, \dots, d\}$, il existe $k_{i,0} \in \{0, 1, \dots, b_i - 1\}$ tel que

$$\{\lambda_i^*\} \in J_{i,0} = \left[\frac{k_{i,0}}{b_i}, \frac{k_{i,0} + 1}{b_i} \right[. \quad (4.30)$$

Par suite, nous définissons

$$I_{i,0} = \{\lambda_i : \langle \lambda_i \rangle = \langle \lambda_i^* \rangle \text{ et } \{\lambda_i\} \in J_{i,0}\}. \quad (4.31)$$

En d'autres termes, pour tout $i \in \{1, \dots, d\}$, nous trouvons

$$I_{i,0} = \begin{cases} \left[\left[\lambda_i^* \right] + \frac{k_{i,0}}{b_i}, \left[\lambda_i^* \right] + \frac{k_{i,0} + 1}{b_i} \right[& \text{si } \lambda_i^* \geq 0, \\ \left[\left[\lambda_i^* \right] - \frac{k_{i,0} + 1}{b_i}, \left[\lambda_i^* \right] - \frac{k_{i,0}}{b_i} \right] & \text{si } \lambda_i^* < 0, \end{cases} \quad (4.32)$$

où $[\cdot]$ représente la fonction partie entière. Par conséquent, nous avons

$$I_0 = \begin{pmatrix} I_{1,0} \\ \vdots \\ I_{d,0} \end{pmatrix}. \quad (4.33)$$

Quelques notations et rappels sont nécessaires. Nous posons,

$$E_0 = \{\theta \in \Theta, \alpha_i = \alpha_i^*, \forall i \in \{1, \dots, d\}, \text{ et } \lambda \in I_0\} = M_0 \times I_0.$$

Notons que $\theta_0 = (M_0, \lambda_0) \in E_0$. Sous [H3] et d'après la Proposition 4.2.1, nous avons

$$K(\theta) - K(\theta_0) = \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2.$$

D'autre part, soit $\theta_* = (M_0, \lambda_*) \in E_0$. D'après la Proposition 4.2.4, nous trouvons

$$K(\theta_*) - K(\theta_0) = \mu_{\theta_0} (\|F(\cdot; \theta) - F(\cdot; \theta_0)\|_2)^2 = 0.$$

Proposition 4.2.5. *Supposons que [H3] soit vérifiée et pour tout $i, j \in \{1, \dots, d\}$, $\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}$ où $a_{ij} \in \mathbb{Z}$, $b_{ij} \in \mathbb{N}^*$ et $a_{ij} \wedge b_{ij} = 1$. Alors, pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta_\varepsilon^0} |K(\theta) - K(\theta_0)| > 0,$$

où $\Theta_\varepsilon^0 = \{\theta \in \Theta, d(\theta, E_0) \geq \varepsilon\}$.

Preuve. Nous distinguons trois situations pour l'événement Θ_ε^0 . Ainsi, notons que

$\Theta_\varepsilon^0 = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, où

$$\Gamma_1 = \{\theta : \exists i_0, j_0 \in \{1, \dots, d\}, |\alpha_{i_0 j_0} - \alpha_{i_0 j_0}^*| \geq \varepsilon\}, \quad (4.34)$$

$$\Gamma_2 = \{\theta : \forall i, j \in \{1, \dots, d\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, \exists l_0 \in \{1, \dots, d\} d(\lambda_{l_0}, I_{l_0,0}) \geq \varepsilon, \langle \lambda_{l_0} \rangle \neq \langle \lambda_{l_0}^* \rangle\}, \quad (4.35)$$

et

$$\Gamma_3 = \{\theta : \forall i, j \in \{1, \dots, d\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, \exists l_0 \in \{1, \dots, d\} d(\lambda_{l_0}, I_{l_0,0}) \geq \varepsilon, \langle \lambda_{l_0} \rangle = \langle \lambda_{l_0}^* \rangle\}. \quad (4.36)$$

L'idée est aussi basée sur l'équation (4.20). Ainsi, le but est de trouver $x_0 \in E$ tel que

$$\|F(x_0; \theta) - F(x_0; \theta_0)\| > 0, \text{ uniformément sur } \Gamma_i, i = 1, 2, 3.$$

- Pour $\theta \in \Gamma_1 \cup \Gamma_2$. Alors, par les mêmes arguments utilisés dans la preuve de la Proposition 4.2.3 pour les sous-ensembles équivalents, nous obtenons

$$\inf_{\theta \in \Gamma_1 \cup \Gamma_2} |K(\theta) - K(\theta_0)| > 0.$$

- Finalement, nous considérons le cas où $\theta \in \Gamma_3$. Pour simplifier les notations, nous supposons que $l_0 = 1$. Ainsi, en particulier, nous avons

$$|\alpha_{1j} - \alpha_{1j}^*| < \varepsilon, \forall j \in \{1, \dots, d\}, d(\lambda_1, I_{1,0}) \geq \varepsilon \text{ et } \langle \lambda_1 \rangle = \langle \lambda_1^* \rangle.$$

Soit $y = (x_1, \dots, x_d)^\tau \in E$. Nous trouvons,

$$\begin{aligned} \|F(y; \theta) - F(y; \theta_0)\|_1 &= \sum_{i=1}^d |f_i(y, \alpha_i, \lambda_i) - f_i(y, \alpha_i^*, \lambda_i^*)|, \\ &\geq |f_1(y, \alpha_1, \lambda_1) - f_1(y, \alpha_1^*, \lambda_1^*)|, \\ &= |\langle \sum_{j=1}^d \alpha_{1j} x_j + \lambda_1 \rangle - \langle \sum_{j=1}^d \alpha_{1j}^* x_j + \lambda_1^* \rangle|. \end{aligned}$$

Soit Θ_0 un compact sous-ensemble de $] -1, 1[^d \times \mathbb{R}$. De plus, nous supposons $\Theta_0 \subset \Theta$, c.à.d. pour tout $\delta = (\beta_1, \dots, \beta_d, \gamma)^\tau \in \Theta_0$ et $\theta = (M, \lambda) \in \Theta$, où $M = (\alpha_{ij})$ et $\lambda = (\lambda_1, \dots, \lambda_d)^\tau \in \mathbb{R}^d$, il existe $0 < A < 1$ et $B > 0$ tels que

$$|\beta_i| \leq |\alpha_{ij}| \leq A \text{ et } |\gamma| \leq |\lambda_j| \leq B, \forall i, j \in \{1, \dots, d\}.$$

D'autre part, nous posons

$$\Gamma_3^0 = \{\theta = (\beta_1, \dots, \beta_d, \gamma)^\tau \in \Theta_0, |\beta_j - \alpha_{1j}^*| < \varepsilon, \forall j \in \{1, \dots, d\}, d(\gamma, I_{1,0}), \text{ et } \langle \gamma \rangle = \langle \lambda_1^* \rangle\}.$$

Sur Γ_3^0 et d'après la preuve de la Proposition 3.2.6, nous obtenons

$$\inf_{\theta \in \Gamma_3^0} |K(\theta) - K(\theta_0)| > 0. \quad \square$$

□

Théorème 4.2.2. *Supposons que [H3] soit vérifiée et pour tout $i, j \in \{1, \dots, d\}$, nous avons $\alpha_{ij}^* = \frac{a_{ij}}{b_{ij}}$, où $a_{ij} \in \mathbb{Z}$, $b_{ij} \in \mathbb{N}^*$ et $a_{ij} \wedge b_{ij} = 1$. Alors,*

$$d(\hat{\theta}_n, E_0) \xrightarrow{p.s.} 0, \text{ où } E_0 = \{M_0\} \times I_0, \text{ avec } I_0 = \begin{pmatrix} I_{1,0} \\ \vdots \\ I_{d,0} \end{pmatrix}.$$

En d'autres termes, $\hat{\alpha}_n$ est fortement consistant tandis que $\hat{\lambda}_n$ converge vers un intervalle de taille $\frac{1}{b}$ contenant λ^* .

Preuve. La conclusion $d(\hat{\theta}_n, E_0) \longrightarrow 0$, presque sûrement, résulte des Propositions 4.2.1 et 4.2.5 par les mêmes arguments utilisés dans la preuve du Théorème 4.2.1. □

4.2.3 La consistance forte de l'estimateur des moindres carrés quand au moins une ligne de la matrice M_0 possède au moins un coefficient irrationnel

Finalement, nous considérons le cas appelé mixte où au moins une ligne de la matrice M_0 possède au moins un coefficient irrationnel et tous les coefficients des autres lignes sont rationnels. Ainsi, il est clair que le résultat de consistance attendu ici n'est autre qu'un mélange des résultats obtenus dans les Théorèmes 4.2.1 et 4.2.2.

Pour simplifier la présentation, nous considérons $d = 2$, la première ligne de la matrice de M_0 possède un coefficient irrationnel, et tous les coefficients de la deuxième ligne de M_0

sont rationnels. Explicitement, nous posons

$$M_0 = (\alpha_{ij}^*) = \begin{pmatrix} a & b \\ \frac{c_1}{d_1} & \frac{c_2}{d_2} \end{pmatrix}, \quad \text{où } a \in \mathbb{R} \setminus \mathbb{Q} \quad \text{et} \quad \lambda_0 = \begin{pmatrix} \lambda_1^* \\ \lambda_2^* \end{pmatrix}. \quad (4.37)$$

De plus, nous définissons

$$\nu = \frac{e}{d_1 d_2}, \quad \text{avec } e = c_1 d_2 \wedge c_2 d_1. \quad (4.38)$$

Nous notons que le numérateur et le dénominateur de ν ne sont pas nécessairement premiers entre eux. Ainsi, nous réécrivons ν_i avec sa forme de fraction irréductible

$$\nu = \frac{p}{q}, \quad \text{où } p \in \mathbb{Z}, q \in \mathbb{N}^* \text{ et } p \wedge q = 1. \quad (4.39)$$

Sans perte de généralité, nous supposons que q est pair. Ainsi, il existe $k_{2,0} \in \{0, \dots, q-1\}$, tel que

$$\lambda_2^* \in \left[\frac{k_{2,0}}{q}, \frac{k_{2,0}+1}{q} \right] = J_{2,0}. \quad (4.40)$$

Par suite, nous définissons

$$I_{2,0} = \{\lambda : \langle \lambda \rangle = \langle \lambda_2^* \rangle, \{\lambda\} \in J_{2,0}\}. \quad (4.41)$$

Finalement, nous notons

$$E'_0 = M_0 \times I'_0, \quad (4.42)$$

où

$$I'_0 = \{(\lambda_1, \lambda_2)^\tau \in \mathbb{R}^2 : \lambda_1 = \lambda_1^* \text{ et } \lambda_2 \in I_{2,0}\}. \quad (4.43)$$

Proposition 4.2.6. *Supposons que [H3] soit vérifiée, M_0 soit définie par (4.37) et q le dénominateur de ν défini par (4.38) et (4.39) soit pair. Alors, pour tout (suffisamment petit) $\varepsilon > 0$, nous avons*

$$\inf_{\theta \in \Theta'_\varepsilon} |K(\theta) - K(\theta_0)| > 0,$$

où

$$\Theta'_\varepsilon = \{\theta : d(\theta, E'_0) \geq \varepsilon\}.$$

Preuve. Nous distinguons cinq situations pour l'événement Θ'_ε . Ainsi, notons que

$$\Theta'_\varepsilon = \Lambda_1 \cup \Lambda_2 \cup \Lambda_3 \cup \Lambda_4 \cup \Lambda_5, \quad \text{où}$$

$$\Lambda_1 = \{\theta : \exists i_0, j_0 \in \{1, 2\}, |\alpha_{i_0 j_0} - \alpha_{i_0 j_0}^*| \geq \varepsilon\}, \quad (4.44)$$

$$\Lambda_2 = \{\theta : \forall i, j \in \{1, 2\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, |\lambda_1 - \lambda_1^*| \geq \varepsilon, \langle \lambda_1 \rangle = \langle \lambda_1^* \rangle\}, \quad (4.45)$$

$$\Lambda_3 = \{ \theta : \forall i, j \in \{1, 2\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, |\lambda_1 - \lambda_1^*| \geq \varepsilon, \langle \lambda_1 \rangle \neq \langle \lambda_1^* \rangle \}, \quad (4.46)$$

$$\Lambda_4 = \{ \theta : \forall i, j \in \{1, 2\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, d(\lambda_2, I_{2,0}) \geq \varepsilon, \langle \lambda_2 \rangle = \langle \lambda_2^* \rangle \}, \quad (4.47)$$

et

$$\Lambda_5 = \{ \theta : \forall i, j \in \{1, 2\}, |\alpha_{ij} - \alpha_{ij}^*| < \varepsilon, d(\lambda_2, I_{2,0}) \geq \varepsilon, \langle \lambda_2 \rangle \neq \langle \lambda_2^* \rangle \}. \quad (4.48)$$

Ainsi, pour tout $i \in \{1, \dots, d\}$, en utilisant les preuves des Propositions 4.2.3 et 4.2.5, il existe $y_{i,0} \in E$ tel que

$$\|F(y_{i,0}; \theta) - F(y_{i,0}; \theta_0)\|_1 > 0, \text{ uniformément sur } \Lambda_i.$$

□

Théorème 4.2.3. *Supposons que [H3] soit vérifiée, M_0 soit définie par (4.37) et q le dénominateur de ν défini par (4.38) et (4.39) soit pair. Alors,*

$$d(\hat{\theta}_n, E'_0) \xrightarrow{p.s.} 0, \text{ où } E_0 = \{M_0\} \times I_0, \text{ avec } I_0 = \begin{pmatrix} \lambda_1^* \\ I_{2,0} \end{pmatrix}.$$

Preuve. La conclusion $d(\hat{\theta}_n, E'_0) \longrightarrow 0$, presque sûrement, résulte des Propositions 4.2.1 et 4.2.6 par les mêmes arguments utilisés dans la preuve du Théorème 4.2.2. □

4.3 Méthode numérique pour calculer $\hat{\theta}_n$

Soient X_0, X_1, \dots, X_n des observations du processus RINVAR(1), défini par (4.3) et (4.4). Rappelons que, pour tout $i \in \{0, 1, \dots, n\}$, nous avons

$$X_i = \begin{pmatrix} X_{1,i} \\ X_{2,i} \\ \vdots \\ X_{d,i} \end{pmatrix}.$$

Notre objectif est de trouver une méthode numérique pour calculer l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\theta \in \Theta} \varphi_n(\theta),$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n (\|X_t - F(X_{t-1}; \theta)\|_2)^2 \quad (4.49)$$

$$= \frac{1}{n} \sum_{t=1}^n |X_{1,t} - f_1(X_{t-1}; \alpha_1, \lambda_1)|^2 + \dots + \frac{1}{n} \sum_{t=1}^n |X_{d,t} - f_d(X_{t-1}; \alpha_d, \lambda_d)|^2. \quad (4.50)$$

Soit Θ_0 un compact sous-ensemble de $] -1, 1[^d \times \mathbb{R}$. De plus, supposons que pour tout $\delta = (\beta_1, \dots, \beta_d, \gamma)^\tau \in \Theta_0$ et $\theta = (M, \lambda) \in \Theta$, où $M = (\alpha_{ij})$ et $\lambda = (\lambda_1, \dots, \lambda_d)^\tau \in \mathbb{R}^d$, il existe $0 < A < 1$ et $B > 0$ tels que

$$|\beta_i| \leq |\alpha_{ij}| \leq A \text{ et } |\gamma| \leq |\lambda_j| \leq B, \quad \forall i, j \in \{1, \dots, d\}.$$

Ainsi, la recherche de $\hat{\theta}_n$ revient à calculer pour tout $i \in \{1, \dots, d\}$

$$(\hat{\alpha}_i, \hat{\lambda}_i) := \arg \min_{(\alpha_i, \lambda_i)} \frac{1}{n} \sum_{t=1}^n |X_{i,t} - f_i(X_{t-1}; \alpha_i, \lambda_i)|^2. \quad (4.51)$$

Pour cela, nous utilisons le même algorithme de recherche dichotomique proposé dans la Section 3.3, pour calculer l'estimateur des moindres carrés des paramètres d'un modèle RINAR(d).

4.4 Analyse des taux de variation et les indices de récolte annuels de la population Suédoise

Dans cette section, nous présentons une application à des données réelles.

Il s'agit des taux de variation (par millier d'habitants) et les indices de récolte (Harvest index) annuels de la population Suédoise entre 1750 et 1849, notées respectivement P_t et h_t , source : Thomas (1940, [60]).

La série P_t a été étudiée dans la Section 2.7. Ainsi, nous rappelons que les observations de cette dernière varient entre -27 et 16 (voir Figure 2.21), sa moyenne empirique est 6.69 et sa variance empirique est 34.559 . De plus, l'ACF et le PACF empiriques de P_t sont données dans la Figure 2.22.

À l'origine, les indices de récolte étaient évalués sur une échelle de demi-point avec un score total de mauvaises récoltes égal à zéro et de récolte surabondante égal à neuf. Pour éviter des observations telles 3.5 , nous raffinons l'échelle de cette série en multipliant par 2 toutes les observations. Ainsi, les indices de récolte annuels h_t varient entre 0 et 18 (voir Figure 4.1). La moyenne empirique de h_t est 10.51 et sa variance empirique est 27.787 . La Figure 4.2 nous donne l'ACF et le PACF empiriques de la série h_t .

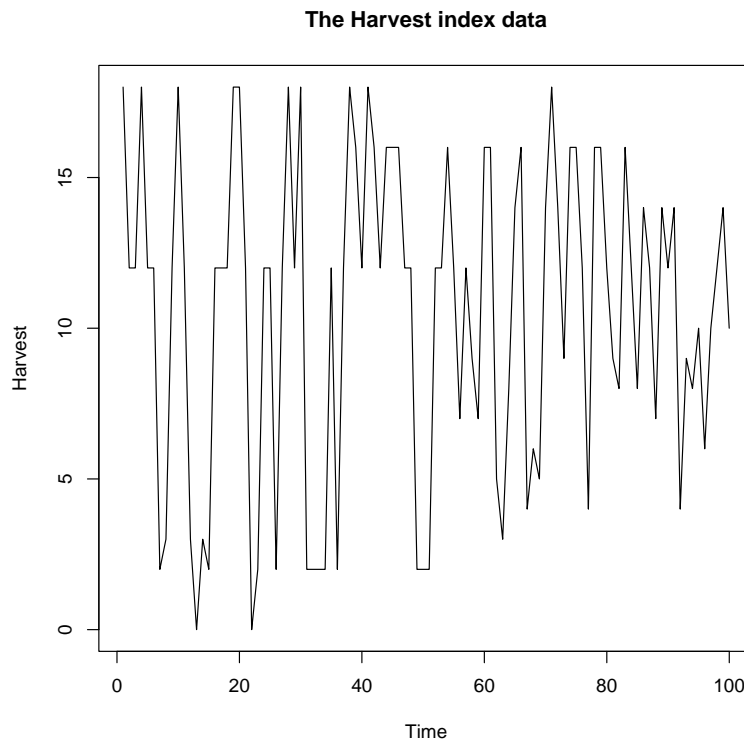
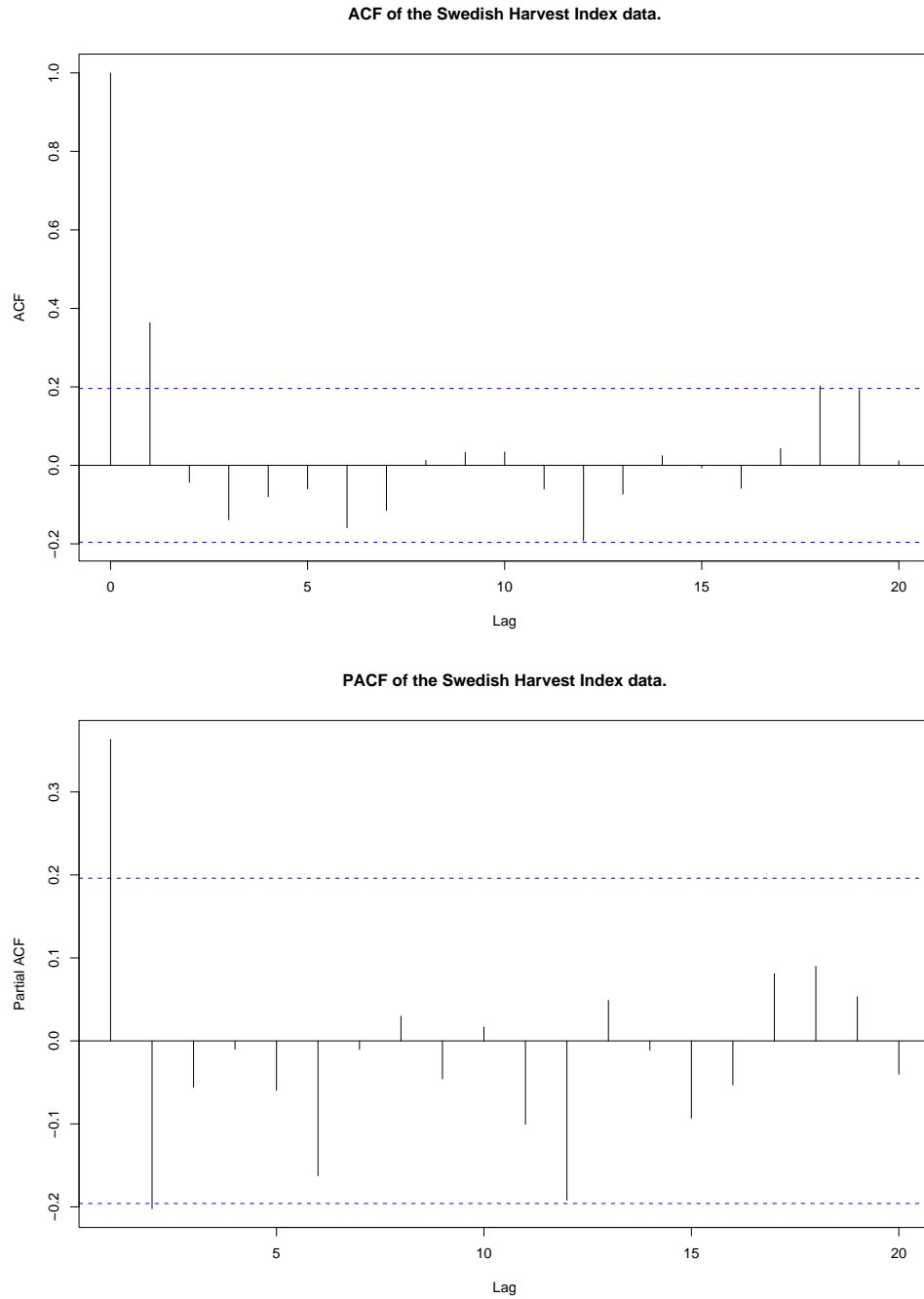


FIG. 4.1 – Les indices de récolte annuels de la population Suédoise, h_t , 1750 – 1849.

FIG. 4.2 – ACF et le PACF empirique de la série h_t .

4.4.1 L'approche de McCleary & Hay

McCleary & Hay (1980, [49]) ont analysé les séries centrées $\bar{P}_t = P_t - \mathbb{E}P_t$ et $\bar{h}_t = h_t - \mathbb{E}h_t$. Pour ces deux séries, les auteurs ont proposé séparément un modèle MA(1) réel.

Ainsi,

$$\bar{h}_t = (1 + \theta_1 B) a_t, \quad \bar{P}_t = (1 + \beta_1 B) b_t, \quad (4.52)$$

où θ_1 et β_1 sont des paramètres réels, (a_t) et (b_t) sont deux suites de bruit centrées.

Pour considérer une interaction possible entre les deux séries, les auteurs ont examiné l'hypothèse suivante : La croissance de la population Suédoise au cours du siècle 1750–1849 pourrait être expliquée presque exclusivement en fonction de la production agricole (les indices de récolte). Ainsi, après l'examen de la fonction de corrélation croisée CCF entre les séries, ils ont proposé le modèle d'interaction suivant :

$$\bar{P}_t = w_0 \bar{h}_{t-1} + N_t, \quad (4.53)$$

où w_0 est un paramètre réel et (N_t) est un bruit centré. En fait, une analyse des résidus suggère le modèle AR(1) pour le bruit.

Par suite, nous avons

$$N_t = (1 - \phi_1 B)^{-1} \xi_t, \quad (4.54)$$

où ϕ_1 est un paramètre réel et ξ_t est un bruit centré.

Ceci mène au modèle suivant

$$\bar{P}_t = w_0 \bar{h}_{t-1} + (1 - \phi_1 B)^{-1} \xi_t. \quad (4.55)$$

Ou d'une manière équivalente, nous trouvons

$$\bar{P}_t = \phi_1 \bar{P}_{t-1} + w_0 (\bar{h}_{t-1} - \phi_1 \bar{h}_{t-2}) + \xi_t. \quad (4.56)$$

Finalement, nous obtenons

$$P_t = \phi_1 P_{t-1} + w_0 (\bar{h}_{t-1} - \phi_1 \bar{h}_{t-2}) + \delta + \xi_t, \quad (4.57)$$

où $\delta = (1 - \phi_1) (\mathbb{E}P_t - w_0 \mathbb{E}h_t)$. Ainsi, la forme du modèle aboutit à une conclusion que la croissance de la population et les indices de récolte sont liés par l'équation

$$P_t = 0.46 P_{t-1} + 0.826 (h_{t-1} - 0.46 h_{t-2}) - 1.075 + \xi_t. \quad (4.58)$$

4.4.2 Ajustement à un modèle RINVAR(1)

Notons que dans la Section 2.7 nous avons proposé, pour analyser la série P_t , le modèle RINAR(1) suivant

$$P_t = \langle 0.469 P_{t-1} + 3.559 \rangle + \varepsilon_t. \quad (4.59)$$

D'autre part, l'ACF et le PACF empiriques de la série h_t (voir Figure 4.2) justifient l'utilisation d'un modèle AR(1) réel (les coefficients du premier ordre sont les seuls significatifs). Ainsi, pour analyser la série des indices de récolte, nous proposons comme pour la série P_t , un modèle RINAR(1). Par suite, en utilisant l'algorithme proposé dans la Section 2.5 pour calculer l'estimateur des moindres carrés des paramètres du modèle RINAR(1), nous obtenons

$$h_t = \langle 0.353 \ h_{t-1} + 7.108 \rangle + \varepsilon_t. \quad (4.60)$$

Nous passons maintenant à une étude multivariée pour analyser les interactions entre les deux séries. D'après l'approche McCleary & Hay, le taux de variation de la population Suédoise à l'instant t dépend de son taux à l'instant $t - 1$ et des indices de récolte aux instants $t - 1$ et $t - 2$. Ainsi, nous considérons le modèle VAR(1) réel suivant

$$\begin{pmatrix} h_t \\ h_{t-1} \\ P_t \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 \\ 1 & 0 & 0 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} h_{t-1} \\ h_{t-2} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} e \\ 0 \\ f \end{pmatrix} + \begin{pmatrix} \varepsilon_{t,1} \\ 0 \\ \varepsilon_{t,2} \end{pmatrix}. \quad (4.61)$$

Donc, en utilisant le logiciel R, nous estimons et analysons les paramètres du modèle (4.61). Les diagnostics de ce dernier indiquent que les coefficients b_1 , c_1 , b_2 et f ne sont pas significatifs. Explicitement, nous trouvons $\hat{b}_1 = -0.170$ (*s.e.*0.109), $\hat{c}_1 = -0.073$ (*s.e.*0.09149), $\hat{b}_2 = -0.04212$ (*s.e.*0.108) et $\hat{f} = -0.42323$ (*s.e.*1.29746).

Ainsi, nous obtenons

$$h_t = a_1 \ h_{t-1} + e + \varepsilon_{t,1}, \quad (4.62)$$

$$P_t = a_2 \ h_{t-1} + c_2 \ P_{t-1} + \varepsilon_{t,2}. \quad (4.63)$$

A partir de cette analyse préliminaire, nous proposons pour les deux séries le modèle RINVAR(1) suivant

$$\begin{pmatrix} h_t \\ P_t \end{pmatrix} = \left\langle \begin{pmatrix} \alpha_{11} & 0 \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} h_{t-1} \\ P_{t-1} \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} \right\rangle + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}. \quad (4.64)$$

Ainsi, en utilisant l'algorithme proposé dans la Section 4.3 pour calculer l'estimateur des moindres carrés du modèle 4.64, nous obtenons

$$\hat{\alpha}_{11} = 0.353, \hat{\alpha}_{21} = 0.380, \hat{\alpha}_{22} = 0.433, \text{ et } \hat{\lambda}_1 = 7.108. \quad (4.65)$$

En particulier, le modèle d'interaction entre les séries est défini par

$$P_t = \langle 0.380 \ h_{t-1} + 0.433 \ P_{t-1} \rangle + \varepsilon_{2,t}. \quad (4.66)$$

4.4.3 Comparaison et commentaires

Notons que la série des taux de variation de la population Suédoise possède des valeurs négatives. Ainsi, la série chronologique bivariée $X_t = (P_t, h_t)^\tau$ ne peut pas être analysée par un modèle MGINAR qui est basé sur l'opérateur d'aminicissement généralisé et défini sur \mathbb{N} .

Afin de comparer les performances de prévision des modèles (4.58) et (4.66), proposés pour analyser l'interaction entre les séries, nous définissons

$$\hat{P}_{t,1} = 0.46 P_{t-1} + 0.826 h_{t-1} - 0.378 h_{t-2} - 1.075, \forall t = 3, \dots, 100 \quad (4.67)$$

et

$$\hat{P}_{t,2} = \langle 0.380 h_{t-1} + 0.433 P_{t-1} \rangle, \forall t = 3, \dots, 100. \quad (4.68)$$

Ici, $\hat{P}_{t,1}$ (resp. $\hat{P}_{t,2}$) représente la prévision à un pas de P_t , basée sur la fonction de régression, provenant du modèle (4.58) (resp. (4.66)).

Notons que $\hat{P}_{t,1}$ est en général une valeur réelle. Le passage au support discret de la série est facilement obtenu en arrondissant à l'entier le plus près, c.à.d.

$$\hat{P}_{t,1} = \langle 0.46 P_{t-1} + 0.826 h_{t-1} - 0.378 h_{t-2} - 1.075 \rangle, \forall t = 3, \dots, 100. \quad (4.69)$$

Notons que le modèle de McCleary & Hay dépend de trois paramètres ϕ_1, w_0 et δ . D'autre part, le modèle RINVAR(1) dépend seulement de deux paramètres. La Figure 4.3 nous montre les vraies valeurs de la série P_t et leurs prévisions $\hat{P}_{t,1}$ et $\hat{P}_{t,2}$. L'erreur quadratique moyenne des prévisions des deux modèles est définie par

$$EQM_i = \frac{1}{20} \sum_{t=3}^{100} \left(P_t - \hat{P}_{t,i} \right)^2, \text{ pour } i = 1, 2.$$

Ainsi, comparée au modèle proposé par McCleary & Hay et défini par (4.58), l'erreur quadratique moyenne des prévisions décroît de 3.7755 à 2.8775.

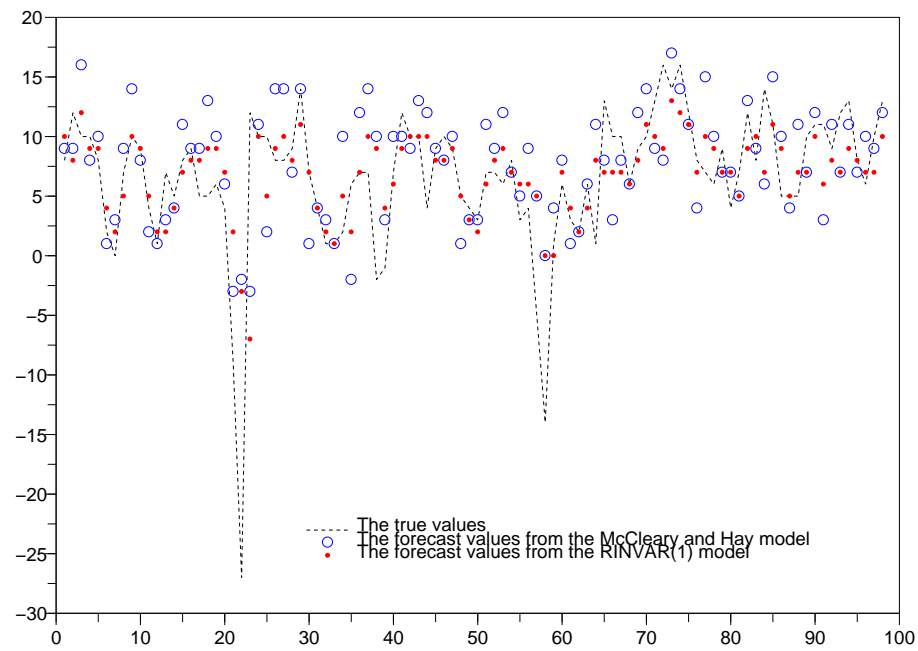


FIG. 4.3 – Les taux de variation de la population Suédoise et leurs prévisions basées sur le modèle de McCleary & Hay et celles basées sur le modèle RINVAR(1).

Chapitre 5

Le modèle RINAR(1) centré

Sommaire

5.1	Quelques définitions	146
5.2	Étude du modèle RINAR(1) centré	148
5.2.1	Stationnarité et ergodicité du processus RINAR(1) centré	149
5.2.2	Estimation du paramètre α	149
5.3	Propriétés du processus RINAR(1) centré	150
5.3.1	Comparaison avec un AR(1) centré	153
5.3.2	Étude de simulation	155
5.4	Méthode numérique pour calculer $\hat{\alpha}_n$	160
5.5	Le modèle RINAR(1) recentré	163

Ce chapitre est consacré à l'étude théorique du modèle RINAR(1) centré, basé sur l'opérateur d'arrondi. Ce dernier représente un cas particulier du modèle RINAR(1) étudié précédemment. Le RINAR(1) centré est utilisé pour analyser des séries chronologiques de moyenne nulle. Dans le premier paragraphe, nous donnons quelques définitions et notations utilisées tout au long du chapitre. Dans le deuxième paragraphe, nous introduisons le modèle RINAR(1) centré, nous donnons les conditions de stationnarité et d'ergodicité du processus et nous étudions la consistance forte de l'estimateur des moindres carrés considéré pour estimer les paramètres du modèle. Dans le troisième paragraphe, nous présentons quelques propriétés du modèle. Par suite, dans le quatrième paragraphe, nous proposons une méthode numérique pour calculer l'estimateur des moindres carrés. Finalement, dans le cinquième paragraphe, nous comparons le modèle RINAR(1) recentré et le processus RINAR(1) centré.

5.1 Quelques définitions

Premièrement, nous définissons la fonction plancher "floor", la partie fractionnaire et la fonction arrondi et nous notons certaines propriétés de chacune.

Définition 5.1.1. *la fonction plancher "floor" pour un nombre réel x , notée $\lfloor x \rfloor$, est une fonction qui renvoie le plus grand nombre entier inférieur ou égal à x . Formellement,*

$$\lfloor x \rfloor = \max \{n \in \mathbb{Z} \mid n \leq x\}.$$

Par exemple : $\lfloor 2.8 \rfloor = 2$, $\lfloor -0.4 \rfloor = -1$, et $\lfloor -2.3 \rfloor = -3$. Pour un réel positif x , $\lfloor x \rfloor$ est la partie entière de x (voir Figure 5.1).

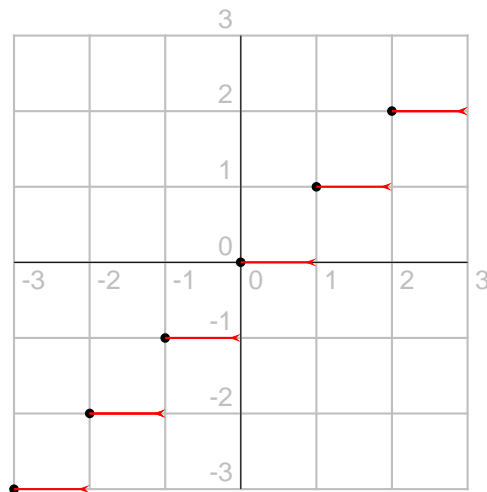


FIG. 5.1 – La fonction "floor"

Définition 5.1.2. *La fonction $x - \lfloor x \rfloor$, aussi écrit comme $x \bmod 1$, ou $\{x\}$, s'appelle la partie fractionnaire de x . Par exemple : $\{3.1\} = 0.1$ et $\{-2.8\} = 0.2$ (c.à.d. si $x < 0$ alors $\{x\} = 1 - \{-x\}$, voir Figure 5.2).*

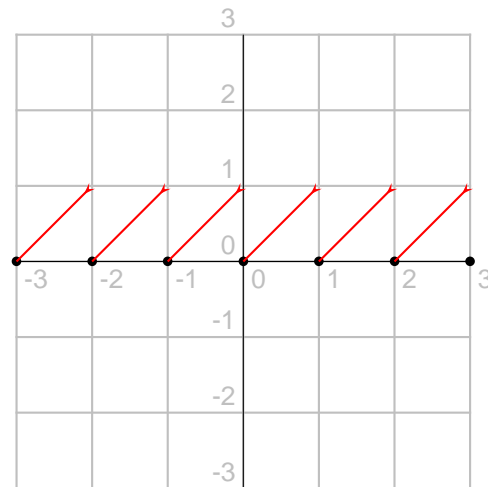


FIG. 5.2 – La fonction partie fractionnaire

Quelques propriétés générales :

1. $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.
2. $\lfloor \lfloor x \rfloor \rfloor = \lfloor x \rfloor$.
3. $0 \leq \{x\} < 1$.
4. $\{\{x\}\} = \{x\}$.
5. $x = \lfloor x \rfloor + \{x\}$.
6. La fonction "floor" n'est pas continue, mais elle est semi-continue supérieur.
7. Soient $k \in \mathbb{Z}$ et $x \in \mathbb{R}$, $\lfloor k + x \rfloor = k + \lfloor x \rfloor$.
8. $\lfloor x \rfloor + \lfloor -x \rfloor = \begin{cases} 0 & \text{si } x \in \mathbb{Z}. \\ -1 & \text{sinon.} \end{cases}$
9. $\{x + \frac{1}{2}\} = \{x - \frac{1}{2}\}$
10. Soient x et y deux réels,
 - $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor + \lfloor \{x\} + \{y\} \rfloor$.
 - $\{x + y\} = \{\{x\} + \{y\}\}$.

Remarque 5.1.1. L'arrondi d'un nombre réel x à l'entier le plus près, noté $\langle x \rangle$ (voir Figure 5.3), peut être exprimé comme

$$\langle x \rangle = \lfloor x + \frac{1}{2} \rfloor.$$

Quelques propriétés générales :

1. $\langle x \rangle = x + \frac{1}{2} - \{x + \frac{1}{2}\}.$
2. $\langle x \rangle = \lfloor x \rfloor + \lfloor \{x\} + \frac{1}{2} \rfloor.$
3. $x - \frac{1}{2} < \langle x \rangle \leq x + \frac{1}{2}.$
4. Soient $k \in \mathbb{Z}$ et $x \in \mathbb{R}$, $\langle k + x \rangle = k + \langle x \rangle.$
5. $\langle \langle x \rangle \rangle = \langle x \rangle.$
6. Soient x et y deux réels, $\langle x + y \rangle = \lfloor x \rfloor + \lfloor y \rfloor + \lfloor \{x\} + \{y\} + \frac{1}{2} \rfloor.$

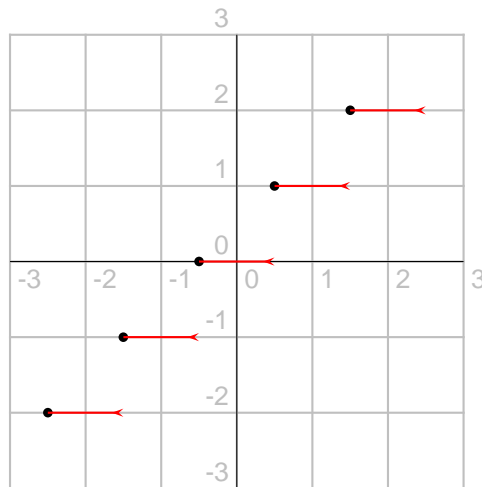


FIG. 5.3 – La fonction arrondi

5.2 Étude du modèle RINAR(1) centré

Nous introduisons un nouveau modèle basé sur l'opérateur d'arrondi, appelé le RINAR(1) centré (pour centered first-order rounded integer-valued autoregressive process). Notons que, ce modèle présente un cas particulier du processus RINAR(1) introduit dans le chapitre 2.

Définition 5.2.1. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Une série de variables aléatoires $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ forme un processus RINAR(1) centré si pour tout $t \in \mathbb{Z}$,

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t, \quad (5.1)$$

où $\langle \cdot \rangle$ représente l'opérateur d'arrondi à l'entier le plus près, (ε_t) est une suite de variables aléatoires i.i.d. centrées à valeurs dans \mathbb{Z} , définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et indépendantes de $\mathcal{F}_{t-1} = \sigma\{X_{t-1}, X_{t-2}, \dots\}$, et α est un paramètre réel.

Notons qu'il est sous-entendu par centré que $\mathbb{E}X_t = 0$, $\forall t \in \mathbb{Z}$. En revanche, à cause de l'opérateur d'arrondi, cette propriété n'est pas une conséquence automatique.

Nous reviendrons sur ce point dans la Section 5.3.

5.2.1 Stationnarité et ergodicité du processus RINAR(1) centré

Le processus (X_t) , défini par (5.1), forme une chaîne de Markov homogène avec un espace d'états $E = \mathbb{Z}$ et une probabilité de transition

$$\pi(x, y) = \mathbb{P}\{\varepsilon_1 = y - \langle \alpha x \rangle\}, \quad x, y \in E. \quad (5.2)$$

En appliquant Proposition 2.2.1, nous avons la proposition suivante qui nous donne les conditions de stationnarité et d'ergodicité du processus RINAR(1) centré.

Proposition 5.2.1. *Supposons que :*

1. *La chaîne de Markov (X_t) est irréductible ;*
2. *pour un certain $k > 1$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;*
3. *$|\alpha| < 1$.*

Alors,

1. *(X_t) possède une unique mesure de probabilité invariante, notée μ . De plus, μ possède un moment d'ordre k .*
2. *Pour tout $x \in E$ et $f \in L^1(\mu)$ nous avons*

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \longrightarrow \mu(f), \quad \mathbb{P}_x \text{ p.s.}$$

où \mathbb{P}_x représente la probabilité conditionnelle $\mathbb{P}(\cdot \mid X_0 = x)$.

5.2.2 Estimation du paramètre α

Nous adaptons la même démarche que dans le Chapitre 2 (Section 2.4). Le problème est plus simple ici puisque le paramètre λ disparaît. De même, il n'y a plus lieu de distinguer les cas " α rationnel" et " α irrationnel".

Soit Θ l'espace des paramètres du modèle RINAR(1) centré. Ce dernier peut être écrit sous la forme suivante

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t = f(X_{t-1}; \alpha) + \varepsilon_t, \quad (5.3)$$

avec $f(x; \alpha) = \langle \alpha x \rangle$, où $x \in \mathbb{Z}$ et $\alpha \in \Theta \subset]-1, 1[$, f est appelée la fonction de régression du modèle.

Soient X_0, X_1, \dots, X_n des observations du processus RINAR(1) centré. Pour l'estimation du paramètre α , nous considérons l'estimateur des moindres carrés défini par

$$\hat{\alpha}_n = \arg \min_{\alpha \in \Theta} \varphi_n(\alpha), \quad (5.4)$$

où

$$\varphi_n(\alpha) = \frac{1}{n} \sum_{t=1}^n (X_t - f(X_{t-1}; \alpha))^2. \quad (5.5)$$

Quelques notations sont nécessaires. Soient α_0 la vraie valeur du paramètre et \mathbb{P}_{α_0} est la mesure de probabilité de la chaîne (X_t) sous le vrai modèle. De plus, toute convergence $\xrightarrow{p.s.}$ signifie une convergence $p.s.$ sous $\mathbb{P}_{\alpha_0, x}$, ce qui se tient indépendamment de l'état initial x . Faisons les hypothèses suivantes.

Hypothèse [H0]

1. Sous \mathbb{P}_{α_0} , la chaîne de Markov (X_t) est irréductible ;
2. pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;
3. α_0 , la vraie valeur de α , appartient à l'intervalle $] -1, 1[$;
4. l'espace des paramètres Θ , est un compact sous-ensemble de $] -1, 1[$.

En particulier, nous trouvons le théorème suivant.

Théorème 5.2.1. *Supposons que [H0] soit vérifiée. Alors, l'estimateur des moindres carrés est fortement consistant, c.à.d.*

$$\hat{\alpha}_n \rightarrow \alpha_0, \quad \mathbb{P}_{\alpha_0} - p.s.$$

5.3 Propriétés du processus RINAR(1) centré

À cause de l'opérateur d'arrondi, le bruit centré et $|\alpha_0| < 1$ ne suffisent pas pour assurer que $\mathbb{E}X_t = 0$. Par ailleurs, nous n'arrivons toujours pas à calculer explicitement le coefficient d'autocorrélation du premier ordre du processus RINAR(1) centré.

Tout d'abord, faisons les hypothèses suivantes.

Hypothèse [A]

1. α_0 , la vraie valeur de α , est un nombre irrationnel ;
2. le bruit ε_t possède une loi symétrique sur $E = \mathbb{Z}$.

Proposition 5.3.1. *Supposons que les hypothèses [A] et [H0] soient vérifiées. Alors,*

1. *La loi stationnaire μ_{α_0} est symétrique sur $E = \mathbb{Z}$, par conséquent $\mathbb{E}X_t = 0$;*
2. *$|\rho(1) - \alpha_0| \leq \frac{1}{2}$.*

De plus, pour tout $j \geq 2$, nous avons

$$|\rho(j) - \alpha_0^j| \leq \frac{1}{2} \sum_{i=0}^{j-1} |\alpha|^i. \quad (5.6)$$

Preuve. Rappelons que le vrai modèle est défini par

$$X_t = \langle \alpha_0 X_{t-1} \rangle + \varepsilon_t = f(X_{t-1}; \alpha_0) + \varepsilon_t,$$

avec $f(x; \alpha_0) = \langle \alpha_0 x \rangle$, $\forall x \in E = \mathbb{Z}$. Notre premier objectif est de montrer que l'unique mesure invariante μ_{α_0} de la chaîne (X_t) est symétrique sur \mathbb{Z} .

Ainsi, pour tout $x \in \mathbb{Z}$, nous avons

$$f(-x; \alpha_0) = \langle -\alpha_0 x \rangle = \lfloor -\alpha_0 x + \frac{1}{2} \rfloor = \lfloor -(\alpha_0 x - \frac{1}{2}) \rfloor.$$

Comme α_0 est un nombre irrationnel, alors $\alpha_0 x - \frac{1}{2} \notin \mathbb{Z}$ et par suite

$$\lfloor -(\alpha_0 x - \frac{1}{2}) \rfloor = -(\lfloor \alpha_0 x - \frac{1}{2} \rfloor + 1).$$

Donc, nous trouvons

$$\begin{aligned} f(-x; \alpha_0) &= -\left(\lfloor \alpha_0 x - \frac{1}{2} \rfloor + 1\right) \\ &= -\left(\alpha_0 x - \frac{1}{2} - \left\{\alpha_0 x - \frac{1}{2}\right\} + 1\right) \\ &= -\left(\alpha_0 x + \frac{1}{2} - \left\{\alpha_0 x - \frac{1}{2}\right\}\right). \end{aligned}$$

Rappelons que, pour tout $a \in \mathbb{R}$, $\{a - \frac{1}{2}\} = \{a + \frac{1}{2}\}$. Ainsi, nous avons

$$\begin{aligned} f(-x; \alpha_0) &= -\left(\alpha_0 x + \frac{1}{2} - \left\{\alpha_0 x + \frac{1}{2}\right\}\right) \\ &= -\lfloor \alpha_0 x + \frac{1}{2} \rfloor \\ &= -\langle \alpha_0 x \rangle \\ &= -f(x; \alpha_0). \end{aligned}$$

Par conséquent, la fonction de régression f est impaire.

De plus, d'après l'hypothèse [A], ε_t et $\eta_t = -\varepsilon_t$ possèdent la même loi.

Maintenant, nous considérons le processus $W = (W_t) = (-X_t)$. Ainsi, nous avons

$$W_t = -f(X_{t-1}; \alpha_0) - \varepsilon_t = f(-X_{t-1}; \alpha_0) + \eta_t = f(W_{t-1}; \alpha_0) + \eta_t.$$

Donc, les lois stationnaires de X_t et W_t coïncident. Par suite, μ_{α_0} est symétrique.

Maintenant, nous calculons $\rho(1)$, le coefficient d'autocorrélation du premier ordre du processus, défini par

$$\rho(1) = \text{corr}(X_t, X_{t+1}) = \frac{\text{cov}(X_t, X_{t+1})}{V(X_t)}, \quad \forall t \in \mathbb{Z}.$$

Comme ε_t est indépendante de $\mathcal{F}_{t-1} = \sigma\{X_{t-1}, X_{t-2}, \dots\}$ et $\mathbb{E}X_t = 0$, il s'ensuit

$$\begin{aligned} \text{cov}(X_t, X_{t+1}) &= \text{cov}(X_t, \langle \alpha_0 X_t \rangle + \varepsilon_{t+1}) \\ &= \mathbb{E}X_t \langle \alpha_0 X_t \rangle \\ &= \alpha_0 \mathbb{E}X_t^2 + \mathbb{E}X_t (\langle \alpha_0 X_t \rangle - \alpha_0 X_t). \end{aligned}$$

Rappelons que $\forall x \in \mathbb{R}, |x - \langle x \rangle| \leq \frac{1}{2}$ et $|X_t| \leq X_t^2, \forall t \in \mathbb{Z}$. Par conséquent,

$$|\rho(1) - \alpha_0| \leq \frac{\mathbb{E}|X_t| |\langle \alpha_0 X_t \rangle - \alpha_0 X_t|}{\mathbb{E}X_t^2} \leq \frac{1}{2} \frac{\mathbb{E}|X_t|}{\mathbb{E}X_t^2} \leq \frac{1}{2}.$$

Supposons que $j = 2$. Nous avons,

$$\begin{aligned} \text{cov}(X_t, X_{t+2}) &= \text{cov}(X_t, \langle \alpha_0 X_{t+1} \rangle + \varepsilon_{t+2}) \\ &= \text{cov}(X_t, \langle \alpha_0 X_{t+1} \rangle) \\ &= \text{cov}(X_t, \alpha_0 X_{t+1} + \langle \alpha_0 X_{t+1} \rangle - \alpha_0 X_{t+1}) \\ &= \alpha_0 \text{cov}(X_t, X_{t+1}) + \mathbb{E}X_t (\langle \alpha_0 X_{t+1} \rangle - \alpha_0 X_{t+1}) \\ &= \alpha_0 (\alpha_0 \mathbb{E}X_t^2 + \mathbb{E}X_t (\langle \alpha_0 X_t \rangle - \alpha_0 X_t)) + \mathbb{E}X_t (\langle \alpha_0 X_{t+1} \rangle - \alpha_0 X_{t+1}). \end{aligned}$$

Alors,

$$\begin{aligned} |\rho(2) - \alpha_0^2| &= \frac{|\alpha_0 \mathbb{E}X_t (\langle \alpha_0 X_t \rangle - \alpha_0 X_t) + \mathbb{E}X_t (\langle \alpha_0 X_{t+1} \rangle - \alpha_0 X_{t+1})|}{\mathbb{E}X_t^2} \\ &\leq \frac{1}{2} (|\alpha_0| + 1). \end{aligned}$$

□

Maintenant, nous considérons le processus RINAR(1) suivant

$$X_t = \langle \alpha_0 X_{t-1} + \lambda_0 \rangle + \varepsilon_t.$$

De plus, nous supposons que les conditions qui assurent la stationnarité et l'ergodicité de ce processus sont vérifiées (voir Section 2.2). Par suite, nous posons $m = \mathbb{E}X_t$, $\forall t \in \mathbb{Z}$.

Rappelons que $|\lambda_0| \leq \frac{1}{2}$ est une condition nécessaire mais non suffisante pour que la moyenne du modèle m soit nulle. D'autre part, si $m \in \mathbb{Z}$ (en particulier $m = 0$), alors les coefficients d'autocorrélation du RINAR(1) (voir Section 2.3), comme ceux du RINAR(1) centré, vérifient l'inégalité suivante

$$|\rho(j) - \alpha_0^j| \leq \frac{1}{2} \sum_{i=0}^{j-1} |\alpha|^i, \quad \forall j \geq 1.$$

5.3.1 Comparaison avec un AR(1) centré

Dans cette section, nous comparons un processus AR(1) centré à valeurs réelles et le modèle RINAR(1) centré ayant le même coefficient de régression.

Soit le processus AR(1) suivant

$$Y_t = \alpha_0 Y_{t-1} + \lambda + \varepsilon_t, \quad (5.7)$$

où (ε_t) est une suite de variables aléatoires i.i.d. centrées et $V(\varepsilon_t) = \sigma^2$.

Ainsi, pour tout $t \in \mathbb{Z}$, nous trouvons

$$\mathbb{E}Y_t = \frac{\lambda}{1 - \alpha_0} \text{ et } V(Y_t) = \frac{\sigma^2}{1 - \alpha_0^2}. \quad (5.8)$$

Rappelons que, sous l'hypothèse $|\alpha_0| < 1$, ce processus peut être écrit comme un modèle MA(∞). Explicitement, nous avons

$$Y_t = \lambda \sum_{j=0}^{\infty} \alpha_0^j + \sum_{j=0}^{\infty} \alpha_0^j \varepsilon_{t-j}. \quad (5.9)$$

Soit le processus RINAR(1) centré suivant

$$X_t = \langle \alpha_0 X_{t-1} \rangle + \varepsilon_t,$$

où (ε_t) est suite de variables aléatoires i.i.d. centrées et à valeurs dans \mathbb{Z} et de variance σ^2 .

Maintenant, nous supposons que les hypothèses [A] et [H0] soient vérifiées.

Ainsi, le processus (X_t) est stationnaire et ergodique. De plus, sa mesure invariante est

symétrique sur E . En d'autres termes, nous obtenons $\mathbb{E}X_t = 0, \forall t \in \mathbb{Z}$.

D'autre part, pour tout $x \in \mathbb{R}$, nous avons

$$\langle x \rangle = \lfloor x + \frac{1}{2} \rfloor = x + \frac{1}{2} - \left\{ x + \frac{1}{2} \right\},$$

où $\{\cdot\}$ représente la partie fractionnaire. Ainsi, le processus RINAR(1) centré, peut être écrit sous la forme suivante

$$X_t = \alpha_0 X_{t-1} + \frac{1}{2} - Z_{t-1} + \varepsilon_t, \quad (5.10)$$

avec $Z_{t-1} = \left\{ \alpha_0 X_{t-1} + \frac{1}{2} \right\}$. Notons que le processus (Z_t) est aussi stationnaire.

Donc, d'après (5.10), nous obtenons

$$(1 - \alpha_0) \mathbb{E}X_t = \frac{1}{2} - \mathbb{E}Z_{t-1} + \mathbb{E}\varepsilon_t$$

Ainsi, comme $\mathbb{E}X_t = 0$ et $\mathbb{E}\varepsilon_t = 0$, nous trouvons

$$\forall t \in \mathbb{Z}, \quad \mathbb{E}Z_t = \frac{1}{2}. \quad (5.11)$$

De plus, pour tout $t \in \mathbb{Z}$, nous avons

$$0 \leq Z_t = \left\{ \alpha_0 X_t + \frac{1}{2} \right\} < 1 \text{ et } Z_t = \begin{cases} \alpha_0 X_t + \frac{1}{2}, & \text{si } 0 \leq \alpha_0 X_t < \frac{1}{2}, \\ \alpha_0 X_t - \frac{1}{2}, & \text{si } \frac{1}{2} \leq \alpha_0 X_t < 1. \end{cases}$$

Notons que la distribution de Z_t est mixte, en d'autres termes elle est continue sur $[0, \frac{1}{2}[$ et $\frac{1}{2}, 1[$ et discrète en $\frac{1}{2}$. Explicitement, nous avons

$$\mathbb{P}(Z_t = \frac{1}{2}) = \mathbb{P}(X_t = 0) + \dots = \mu_{\alpha_0}(\{0\}) + \dots \neq 0.$$

En itérant l'équation (5.10), nous obtenons

$$\begin{aligned} X_t &= \varepsilon_t + \alpha_0 \varepsilon_{t-1} + \dots + \alpha_0^k \varepsilon_{t-k} \\ &\quad + \frac{1}{2} + \frac{\alpha_0}{2} + \dots + \frac{\alpha_0^k}{2} \\ &\quad - Z_{t-1} - \alpha_0 Z_{t-2} - \dots - \alpha_0^k Z_{t-k-1}. \end{aligned}$$

Comme $|\alpha_0| < 1$ et (X_t) est stationnaire (par conséquent $\forall t \in \mathbb{Z}, \mathbb{E}X_t^2$ est une constante), nous en déduisons qu'au sens de la convergence dans L^2 (par le critère de Cauchy), nous avons

$$X_t = \frac{1}{2} \sum_{j=0}^{\infty} \alpha_0^j + \sum_{j=0}^{\infty} \alpha_0^j \varepsilon_{t-j} - \sum_{j=0}^{\infty} \alpha_0^j Z_{t-j-1} \quad (5.12)$$

$$= \sum_{j=0}^{\infty} \alpha_0^j \varepsilon_{t-j} + \sum_{j=0}^{\infty} \alpha_0^j \left(\frac{1}{2} - Z_{t-j-1} \right). \quad (5.13)$$

En effet, nous trouvons

$$\mathbb{E} \left[X_t - \frac{1}{2} \sum_{j=0}^K \alpha_0^j - \sum_{j=0}^K \alpha_0^j \varepsilon_{t-j} + \sum_{j=0}^K \alpha_0^j Z_{t-j-1} \right]^2 = \alpha_0^{2K+2} \mathbb{E} X_{t-(k+1)}^2 \rightarrow 0, \text{ lorsque } k \rightarrow \infty.$$

Maintenant, nous comparons les deux processus (X_t) et (Y_t) . Nous distinguons deux cas :

– Si $\lambda = 0$, alors $\mathbb{E}Y_t = \mathbb{E}X_t = 0$ et

$$|X_t - Y_t| = \left| \sum_{j=0}^{\infty} \alpha_0^j \left(\frac{1}{2} - Z_{t-j-1} \right) \right| \quad (5.14)$$

$$\leq \frac{1}{2(1 - |\alpha_0|)}. \quad (5.15)$$

– Si $\lambda = \frac{1}{2}$, alors $\mathbb{E}Y_t = \frac{1}{2(1 - \alpha_0)}$ et

$$|X_t - Y_t| = \left| \sum_{j=0}^{\infty} \alpha_0^j Z_{t-j-1} \right| \quad (5.16)$$

$$\leq \frac{1}{(1 - |\alpha_0|)}. \quad (5.17)$$

D'autre part, d'après (5.10), nous avons

$$\begin{aligned} V(X_t) &= V(\alpha_0 X_t - Z_t) + \sigma^2, \\ &= \alpha_0^2 V(X_t) + V(Z_t) - 2\alpha_0 \text{cov}(X_t, Z_t) + \sigma^2, \\ &= \alpha_0^2 V(X_t) + V(Z_t) - 2\alpha_0 \mathbb{E}X_t Z_t + \sigma^2. \end{aligned}$$

Comme $|\alpha_0| < 1$, alors

$$V(X_t) = \frac{\sigma^2}{1 - \alpha_0^2} + \frac{V(Z_t) - 2\alpha_0 \mathbb{E}X_t Z_t}{1 - \alpha_0^2}. \quad (5.18)$$

Ainsi, nous obtenons

$$|V(X_t) - V(Y_t)| = \frac{|V(Z_t) - 2\alpha_0 \mathbb{E}X_t Z_t|}{1 - \alpha_0^2}. \quad (5.19)$$

5.3.2 Étude de simulation

Le RINAR(1) centré est inspiré du modèle AR(1) réel. Rappelons que les coefficients d'autocorrélation théoriques d'un processus AR(1), ayant α_0 comme coefficient de régression, sont donnés par

$$\rho(k) = \alpha_0^k, \quad \forall k \geq 1. \quad (5.20)$$

Ici, notre but est de comparer les fonctions d'autocorrélation d'un RINAR(1) centré et un AR(1) possédant le même coefficient de régression α_0 . Pour cela, nous procédons par

simulation. Pour obtenir un bruit centré à valeurs dans \mathbb{Z} , nous considérons deux suites ξ_t et η_t indépendantes de variables i.i.d. suivant une loi de Poisson de paramètre μ , et nous posons $\varepsilon_t = \xi_t - \eta_t$. Ainsi, nous distinguons deux cas : $\mu \gg 1$ et $0 < \mu < \frac{1}{2}$ correspondant respectivement à un grand et petit bruit.

Cette notion du bruit est relative à sa variance $V(\varepsilon_t) = 2\mu$. Plus précisément, d'après l'inégalité de Bienaymé-Tchebychev, nous avons

$$\mathbb{P}\left[|\varepsilon_t| \leq k \sqrt{2\mu}\right] \geq 1 - \frac{1}{k^2}, \quad \forall k > 0.$$

Par exemple, pour $\mu = \frac{1}{2}$ et $k = 10$, nous avons $\mathbb{P}[\varepsilon_t \in \{-10, \dots, 0, \dots, 10\}] \geq 0.99$.

- Cas 1 : Nous simulons 1000 observations du processus RINAR(1) centré, avec $\alpha = 0.83$ et $\mu = 9.453$ (grand bruit).

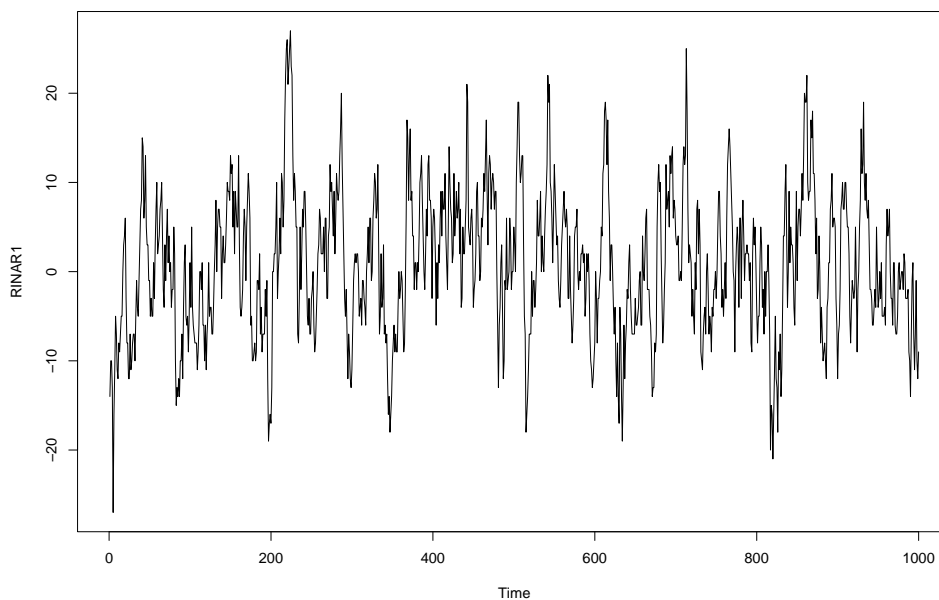


FIG. 5.4 – Les observations simulées du RINAR(1) centré, avec $\alpha_0 = 0.83$ et $\mu = 9.453$

Les observations ci-dessus montrent que les valeurs oscillent autour de 0, comme pour un processus AR(1) centré réel. Par suite, nous comparons les coefficients d'autocorrélation théoriques d'un AR(1) standard (avec $\alpha_0 = 0.83$), définis par (5.20), et les ACF empiriques de la série simulée.

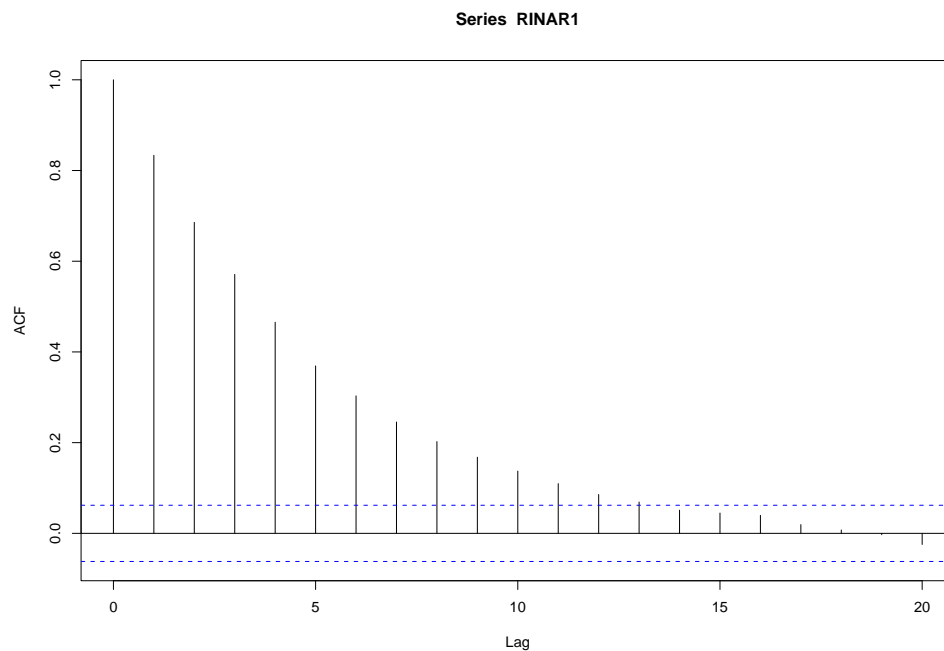
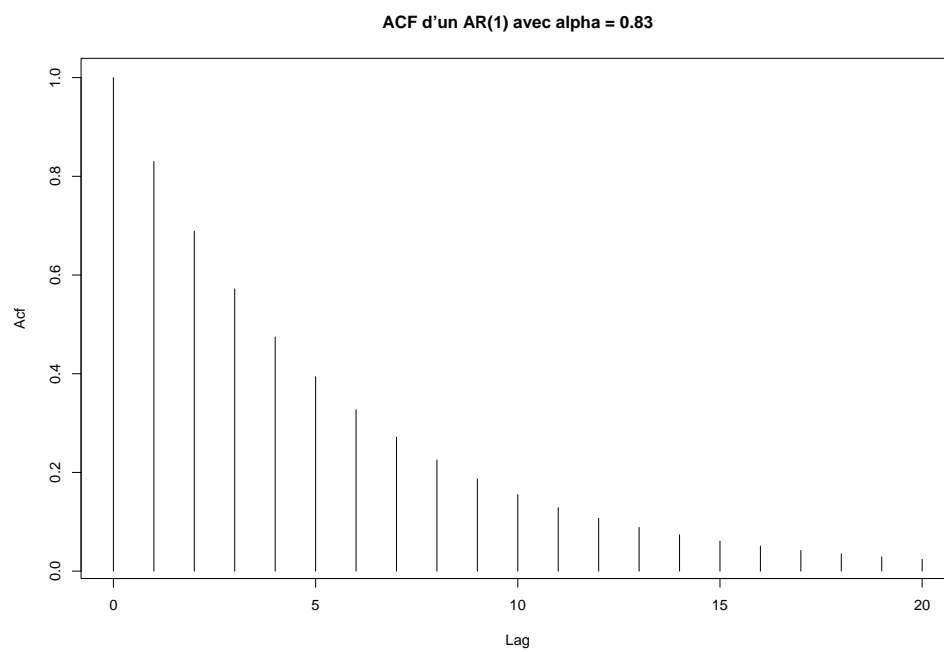


FIG. 5.5 – ACF de la série simulée

FIG. 5.6 – ACF théorique d'un AR(1) avec $\alpha_0 = 0.83$

Nous remarquons que les ACF empiriques de la série simulée possèdent la même décroissance exponentielle que les ACF théoriques du processus AR(1) ayant le même

coefficient de régression α_0 .

- Cas 2 : Nous simulons 1000 observations du processus RINAR(1) centré, avec $\alpha_0 = 0.83$ et $\mu = 0.1274$ (petit bruit).

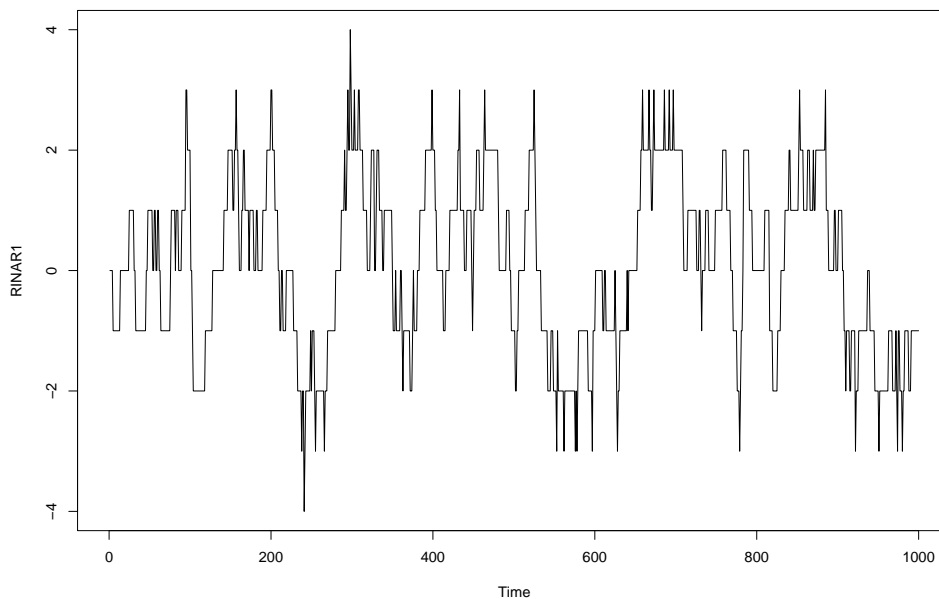


FIG. 5.7 – Les observations simulées du RINAR(1) centré, avec $\alpha_0 = 0.83$ et $\mu = 0.1274$

Les observations ci-dessus ont un comportement particulier. Comme le bruit est petit et à cause de l'arrondi, les valeurs restent souvent invariantes sur une longue période de temps.

Ainsi, rappelons que $\mathbb{E}\varepsilon_t = 0$ et $V(\varepsilon_t) = 2\mu \approx 0.25$. Soit $a > 0$, d'après l'inégalité de Bienaymé-Tchebychev, nous avons

$$\mathbb{P}\left(-a\sqrt{2\mu} < \varepsilon_t < a\sqrt{2\mu}\right) \geq 1 - \frac{1}{a^2}.$$

- Pour $a = 2$, nous obtenons

$$\mathbb{P}(\varepsilon_t = 0) \geq 0.75.$$

- Pour $a = 4$, nous trouvons

$$\mathbb{P}(\varepsilon_t \in \{-1, 0, 1\}) \geq 0.9375.$$

De plus, pour tout $x, y \in \mathbb{Z}$, nous avons

$$\pi(x, y) = \mathbb{P}(X_{t+1} = y \mid X_t = x) = \mathbb{P}(\varepsilon_1 = y - \langle \alpha_0 x \rangle), \forall t \in \mathbb{Z}.$$

Donc,

$$\mathbb{P}(X_{t+1} = 0 \mid X_t = 0) = \mathbb{P}(\varepsilon_1 = 0) \geq 0.75.$$

Comme $\alpha_0 = 0.83 > \frac{1}{2}$, nous avons

$$\mathbb{P}(X_{t+1} = -1 \mid X_t = -1) = \mathbb{P}(X_{t+1} = 1 \mid X_t = 1) = \mathbb{P}(\varepsilon_1 = 0) \geq 0.75.$$

De plus, $2\alpha_0 = 1.66$ et par conséquent $\langle 2\alpha_0 \rangle = 2$ et $\langle -2\alpha_0 \rangle = -2$, nous obtenons

$$\mathbb{P}(X_{t+1} = -2 \mid X_t = -2) = \mathbb{P}(X_{t+1} = 2 \mid X_t = 2) = \mathbb{P}(\varepsilon_1 = 0) \geq 0.75.$$

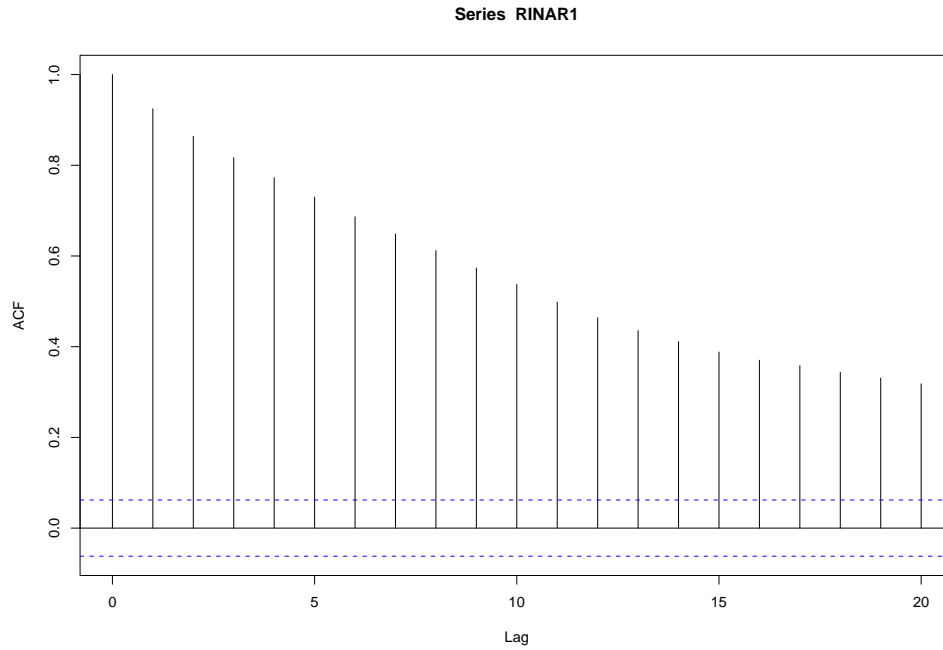


FIG. 5.8 – ACF de la série simulée

La décroissance des ACF empiriques de la série simulée est plutôt affine, elle ressemble plus à celle des ACF théorique d'un AR(1) avec un coefficient de régression assez élevé (par exemple $\alpha_0 = 0.95$).

Finalement, nous constatons que pour le cas où le bruit est grand le coefficient d'auto-corrélation du premier ordre empirique, $\hat{\rho}(1)$, est une bonne estimation pour le paramètre de régression α_0 .

En revanche, pour le cas où (ε_t) est petit, la dépendance entre les variables est plus forte. Par conséquent, nous remarquons que $\hat{\rho}(1)$ est éloigné de α_0 .

Ainsi, les résultats obtenus pour les deux cas sont compatibles avec la conclusion 2 de la Proposition 5.3.1.

5.4 Méthode numérique pour calculer $\hat{\alpha}_n$

Dans cette section, nous supposons que [A] et [H0] soient vérifiées. Quelques rappels sont nécessaires. Soient X_0, X_1, \dots, X_n des observations provenant du processus RINAR(1) centré,

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t.$$

Pour estimer α , nous considérons l'estimateur des moindres carrés $\hat{\alpha}_n$, défini par

$$\hat{\alpha}_n = \arg \min_{\alpha \in \Theta} \mathbb{P}_n g_\alpha,$$

où \mathbb{P}_n est la mesure empirique engendrée par Y_1, Y_2, \dots, Y_n , avec $Y_t = (X_{t-1}, X_t)$, et g_α est définie par

$$g_\alpha((x, y)) = (y - \langle \alpha x \rangle)^2, \quad \forall (x, y) \in E^2 = \mathbb{Z}^2.$$

À cause de l'arrondi, la fonction de contraste g_α est irrégulière. Par conséquent, les méthodes numériques classiques pour calculer $\hat{\alpha}_n$, comme la descente de gradient, ne sont pas valables.

En revanche, sous les hypothèses [A] et [H0], nous avons

$$\left\{ \begin{array}{l} -1 < \alpha < 1, \\ \text{et} \\ \rho(1) - \frac{1}{2} \leq \alpha \leq \rho(1) + \frac{1}{2}. \end{array} \right.$$

Pour calculer $\hat{\alpha}_n$, nous proposons une simple méthode probabiliste. Il s'agit de choisir uniformément K valeurs (par exemple $K \geq 1000$) $\{\alpha_k, 1 \leq k \leq K\}$, sur l'intervalle obtenu par intersection de $] -1, 1[$ et $[\rho(1) - \frac{1}{2}, \rho(1) + \frac{1}{2}]$. Puis, nous calculons $\mathbb{P}_n g_{\alpha_k}$ pour chacune de ces K valeurs. Soit α_{k_0} tel que $\mathbb{P}_n g_{\alpha_{k_0}} = \min \{\mathbb{P}_n g_{\alpha_k}, 1 \leq k \leq K\}$.

Finalement, nous obtenons $\hat{\alpha}_n = \alpha_{k_0}$.

Validation par une étude de simulation

Notre objectif est d'illustrer les performances de la méthode choisie pour calculer $\hat{\alpha}_n$. Pour cela, nous simulons 1000 observations du processus RINAR(1) centré, avec α_0 fixé. Pour obtenir un bruit centré à valeurs dans \mathbb{Z} , nous considérons deux suites ξ_t et η_t indépendantes de variables i.i.d. suivant une loi de Poisson de paramètre μ , et nous posons $\varepsilon_t = \xi_t - \eta_t$.

Puis, en utilisant la méthode indiquée précédemment avec $K = 1000$, nous calculons $\hat{\alpha}_n$. Après 500 répétitions indépendantes, nous calculons la moyenne et l'écart-type de la suite des estimations obtenues $\{\hat{\alpha}_{n,i}, 1 \leq i \leq 500\}$, et dessinons leurs histogrammes.

Précédemment, nous avons distingué deux cas relativement au bruit du modèle.

Ainsi, si ce dernier est grand le coefficient d'autocorrélation du premier ordre est proche de α_0 . Par conséquent, $\hat{\alpha}_n$ doit être proche de $\hat{\rho}(1)$.

En revanche, si le bruit est petit $\rho(1)$ est éloigné de α_0 . Ainsi, $\hat{\alpha}_n$ s'écarte de $\hat{\rho}(1)$.

Ici, nous restons fidèle au jeu de paramètres considéré dans l'étude de simulation de la section 5.3.2.

- Cas du grand bruit : $\alpha_0 = 0.83$ et $\mu = 9.453$.

La figure ci-dessous montre la répartition des estimations obtenues. La moyenne des estimations, notée $\bar{\alpha}_n$, est égale 0.823 et l'écart-type est 0.016.

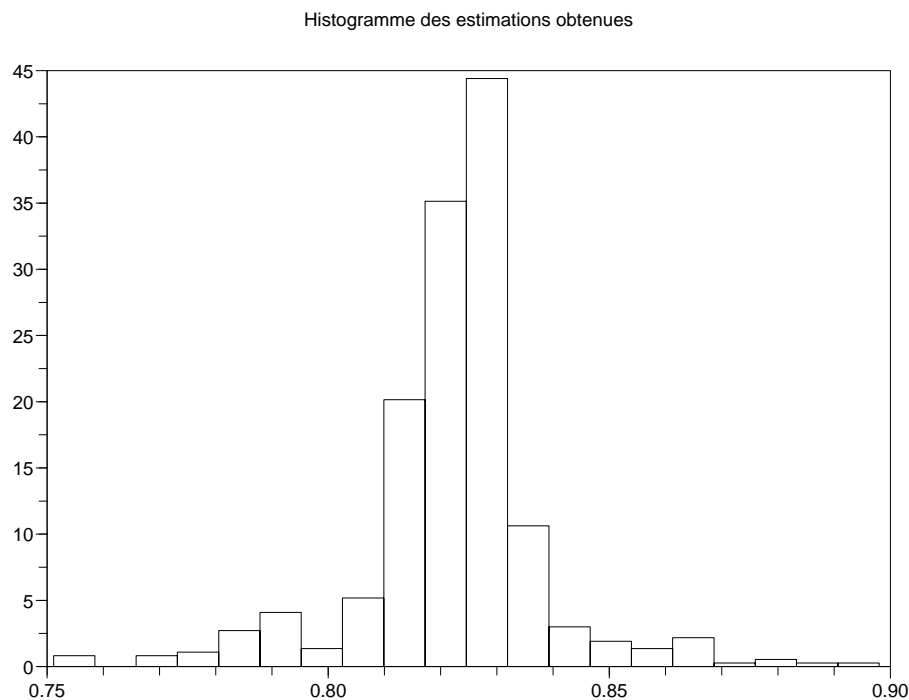


FIG. 5.9 – Histogramme des estimations de $\alpha_0 = 0.83$ (avec grand bruit)

D'autre part, notons que la moyenne des coefficients d'autocorrélations du premier ordre, notée $\bar{\hat{\rho}}(1)$, est égale à 0.821 et l'écart-type est 0.019. Donc, avec un grand bruit, nous pouvons dire que $\hat{\rho}(1)$ est un bon estimateur de α_0 , comme pour un

AR(1) réel.

- Cas du petit bruit : $\alpha_0 = 0.83$ et $\mu = 0.1274$.

Les figures suivantes montrent respectivement la répartition des coefficients d'auto-corrélation empiriques du premier ordre et des estimations obtenues.

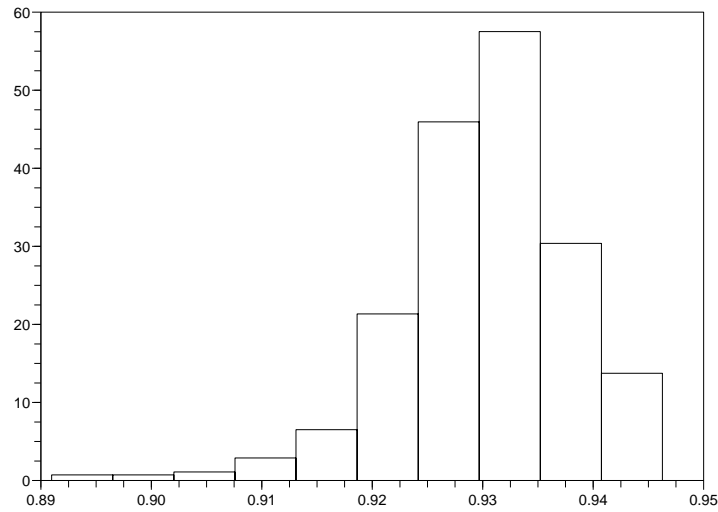


FIG. 5.10 – Histogramme des $\hat{\rho}_k(1)$, $1 \leq k \leq 500$

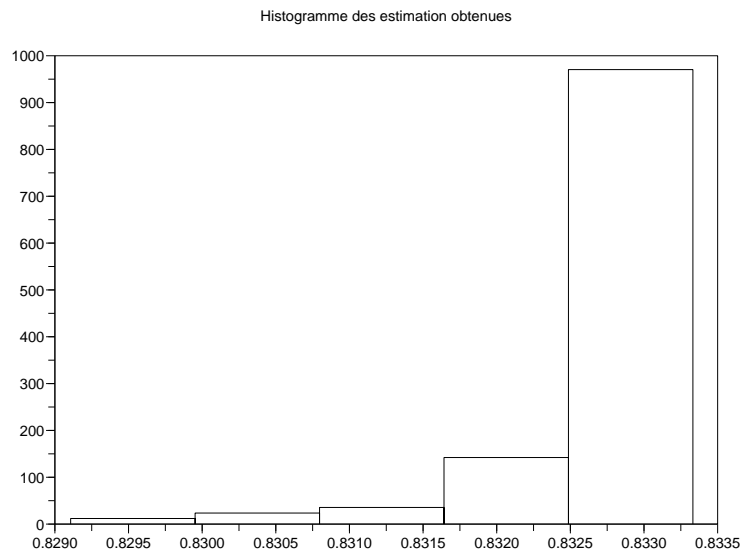


FIG. 5.11 – Histogramme des estimations de $\alpha_0 = 0.83$ (avec petit bruit)

Notons que $\bar{\rho}(1) = 0.93$ et l'écart-type est de 0.008. Par conséquent, dans le cas du petit bruit, il est clair que le coefficient d'autocorrélation du premier ordre empirique n'est pas un bon estimateur de α_0 . En revanche, nous avons $\bar{\alpha}_n = 0.832$ et l'écart-type des estimations obtenues est égale 0.001.

5.5 Le modèle RINAR(1) recentré

Soient X_1, \dots, X_n des observations à valeurs entières. Le procédé habituel de centrage de cette série chronologique semble inapproprié pour une modélisation à valeurs entières. En effet, la moyenne empirique $\hat{m} = \bar{X}_n$ est généralement non entière. Par conséquent, le processus $Y_t = X_t - \hat{m}$ ne rentre pas dans le cadre des modèles considérés.

Dès lors que le processus RINAR(1),

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t,$$

a été légitimé pour modéliser la série chronologique X_1, \dots, X_n (par exemple, en testant si les résidus peuvent être considérés comme un bruit blanc), une question naturelle se pose : Pouvons-nous, comme dans le cas d'un AR(1) réel, nous ramener de manière systématique à l'étude d'une série centrée ?

Autrement dit, est-ce que le processus RINAR(1) recentré, $Y_t = X_t - m$ où $m = \mathbb{E}X_t$, possède la même forme qu'un processus RINAR(1) centré ?

Plus explicitement, sous l'hypothèse $m \in \mathbb{Z}^*$, avons-nous

$$Y_t \stackrel{?}{=} \langle \alpha Y_{t-1} \rangle + \varepsilon_t.$$

Ainsi, nous avons

$$X_t - m = \langle \alpha (X_{t-1} - m) + \alpha m + \lambda \rangle - m + \varepsilon_t.$$

Comme $m \in \mathbb{Z}^*$, nous trouvons

$$Y_t = \langle \alpha Y_{t-1} + c \rangle + \varepsilon_t, \tag{5.21}$$

où $c = \lambda - m(1 - \alpha)$. D'après la Proposition 2.3.1, nous avons

$$|c| \leq \frac{1}{2}. \tag{5.22}$$

Par définition, le processus (Y_t) est centré. De plus, d'après la preuve de la Proposition 2.3.2, les coefficients d'autocorrélation de ce dernier possèdent les mêmes encadrements que

le processus RINAR(1) centré. En particulier, nous avons

$$|\rho(1) - \alpha| \leq \frac{1}{2}.$$

Finalement, sous l'hypothèse $m \in \mathbb{Z}$, le processus RINAR(1) recentré présente quelques avantages. En effet, si les observations de la série chronologique en question sont assez élevées, le processus (Y_t) permet de traiter des observations à faibles variations et oscillant autour de 0. D'autre part, le paramètre de la constante c du processus RINAR(1) recentré varie entre $-\frac{1}{2}$ et $\frac{1}{2}$. Ainsi, nous proposons $\hat{\theta}_0 = (\hat{\alpha}_0, \hat{c}_0) = (\hat{\rho}(1), 0)$ comme point de départ de l'algorithme utilisé pour calculer l'estimateur des moindres carrés (voir Section 2.5), considéré pour estimer les paramètres du modèle.

Maintenant, nous supposons que $m \notin \mathbb{Z}$ et les observations de la série chronologique soient assez élevées. Ainsi, les observations du processus (Y_t) ne sont plus entières. En revanche, nous pouvons envisager d'autres processus à valeurs entières dont la moyenne est faible mais pas forcément nulle, les observations varient autour de 0, et les coefficients d'autocorrélation possèdent les mêmes encadrements que ceux d'un processus RINAR(1). À titre d'exemple, nous pensons au processus $M_t = X_t - me$ où me est la médiane des observations de la série. De plus, la médiane est bon estimateur de la moyenne ayant en général la même nature entière de la série.

Autrement, nous pouvons considérer le processus suivant

$$W_t = X_t - \langle m \rangle.$$

Notons que ce dernier est à valeurs entières et $|\mathbb{E}W_t| = |m'| = |m - \langle m \rangle| \leq \frac{1}{2}$.

D'autre part, le processus (W_t) peut être écrit sous la forme suivante

$$W_t = \langle \alpha W_{t-1} + c_1 \rangle + \varepsilon_t, \quad (5.23)$$

où $c_1 = \lambda - \langle m \rangle (1 - \alpha)$. De plus, comme $-1 < \alpha < 1$, nous avons

$$\begin{aligned} |c_1| &= |\lambda - m(1 - \alpha) + m(1 - \alpha) - \langle m \rangle(1 - \alpha)| \\ &\leq |\lambda - m(1 - \alpha)| + |m(1 - \alpha) - \langle m \rangle(1 - \alpha)| \\ &\leq \frac{1}{2} + \frac{1}{2}(1 - \alpha) < \frac{3}{2}. \end{aligned}$$

Les coefficients d'autocorrélation du processus (W_t) possèdent les mêmes encadrements que ceux d'un modèle RINAR(1) avec une moyenne non entière (voir Proposition 2.3.3).

Chapitre 6

Le modèle PRINAR(1)

Sommaire

6.1	Introduction	166
6.2	Étude théorique du processus PRINAR(1)	167
6.3	Propriétés du processus PRINAR(1)	168
6.3.1	Le processus PRINAR(1) recentré	171
6.3.2	Étude de simulation	172
6.4	Support de la loi stationnaire	177

Dans ce chapitre, nous proposons un processus autorégressif basé sur l'opérateur d'arrondi pour analyser des séries chronologiques à valeurs entières positives.

Dans le premier paragraphe, nous introduisons le modèle PRINAR(1). Dans le deuxième paragraphe, nous donnons le résultat principal concernant la consistance forte de l'estimateur des moindres carrés considéré pour estimer les paramètres du modèle. Dans le troisième paragraphe, nous étudions les propriétés générales du modèle. Par suite, dans le quatrième paragraphe, nous donnons les conditions nécessaires pour que la loi stationnaire du processus possède un support fini.

6.1 Introduction

Tout d'abord, notons que nous gardons les mêmes notations et définitions que dans le chapitre précédent, notamment la fonction plancher "floor", la fonction partie fractionnaire et la fonction arrondi. Soit le modèle suivant

$$X_t = \langle \alpha X_{t-1} \rangle + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (6.1)$$

où (ε_t) une suite de variables aléatoires i.i.d. à valeurs dans \mathbb{N} , avec $\mathbb{E}\varepsilon_t = \mu$ (par exemple, $\varepsilon_t \rightsquigarrow Po(\mu)$), et α est un paramètre réel positif. Nous appelons ce modèle PRINAR(1) (pour positive first-order rounded integer-valued autoregressive).

Notons que, sous l'hypothèse $\alpha \geq 0$, la positivité du processus est assurée, c.à.d.

$$\forall t \in \mathbb{Z}, X_t \geq 0 \implies X_{t+1} \geq 0.$$

Alors, à première vue le processus PRINAR(1) ne peut pas produire des autocorrélations négatives.

D'autre part, la prévision à un pas, basée sur l'espérance conditionnelle, est donnée par

$$\hat{X}_{T+1} = \mathbb{E}(X_{T+1} \mid \mathcal{F}_T) = \langle \alpha X_T \rangle + \mu, \quad (6.2)$$

où $\mathcal{F}_T = \sigma\{X_T, X_{T-1}, \dots\}$. Notons que \hat{X}_{T+1} , pour un RINAR(1) centré par exemple, est directement une valeur entière. Ici, \hat{X}_{T+1} est plutôt une valeur réelle.

Le passage au support discret de la série est facilement obtenu en arrondissant à l'entier le plus près, c.à.d.

$$\hat{X}_{T+1} = \langle \langle \alpha X_T \rangle + \mu \rangle = \langle \alpha X_T \rangle + \langle \mu \rangle. \quad (6.3)$$

Remarquons que, pour $0 < \mu < \frac{1}{2}$, nous obtenons

$$\hat{X}_{T+1} = \langle \alpha X_T \rangle. \quad (6.4)$$

Donc dans ce cas, μ n'a aucun effet sur la prévision à un pas, basée sur l'espérance conditionnelle.

Ainsi, malgré que sa moyenne est non nulle, le modèle PRINAR(1) avec un petit bruit possède une formule de prévision à un pas identiques à celles d'un RINAR(1) centré.

Notons que, sous l'hypothèse $\Theta \subset [0, 1[\times \mathbb{N}^*$, en particulier μ est entier, le processus PRINAR(1) défini par (6.1) peut être écrit sous la forme suivante

$$X_t = \langle \alpha X_{t-1} + \mu \rangle + \varepsilon'_t, \quad t \in \mathbb{Z}, \quad (6.5)$$

où $\varepsilon'_t = \varepsilon_t - \mu$ est bruit centré et à valeurs dans \mathbb{Z} .

Par suite, il est clair que le modèle présent représente un cas particulier du modèle RINAR(1). De plus, la prévision à un pas donnée par (6.2) est directement une valeur entière.

6.2 Étude théorique du processus PRINAR(1)

L'étude de la stationnarité, l'ergodicité, et la consistance de l'estimateur des moindres carrés du processus PRINAR(1) est similaire à celle du modèle RINAR(1) du Chapitre 2. Ici, nous introduisons quelques notations et annonçons le résultat principal.

PRINAR(1) peut être écrit sous la forme suivante

$$X_t = \langle \alpha X_{t-1} \rangle + \mu + \varepsilon'_t = f(X_{t-1}; \theta) + \varepsilon'_t, \quad (6.6)$$

où $\varepsilon'_t = \varepsilon_t - \mu$ ($\mathbb{E}\varepsilon'_t = 0$) et $f(x; \theta) = \langle \alpha x \rangle + \mu$, $\forall x \in \mathbb{N}$ avec $\theta = (\alpha, \mu) \in \Theta \subset \mathbb{R}^+ \times \mathbb{R}^{+*}$. Soient $\theta_0 = (\alpha_0, \mu_0)$ la vraie valeur du paramètre et \mathbb{P}_{θ_0} est la mesure de probabilité de la chaîne (X_t) sous le vrai modèle. De plus, toute convergence $\xrightarrow{p.s.}$ signifie une convergence *p.s.* sous $\mathbb{P}_{\theta_0, x}$, ce qui se tient indépendamment de l'état initial x .

Le fait que la constante μ soit hors de l'opérateur d'arrondi présente un avantage par rapport au modèle RINAR(1) du chapitre 2. En particulier, le processus PRINAR(1) n'a pas de problème d'identifiabilité.

Ainsi, pour tout $\theta \in \Theta$, nous obtenons

$$f(x; \theta) = f(x; \theta_0), \quad \forall x \in \mathbb{Z} \iff \theta = \theta_0.$$

Soient X_0, X_1, \dots, X_n des observations du processus PRINAR(1). Pour l'estimation du paramètre θ , nous considérons l'estimateur des moindres carrés défini par

$$\hat{\theta}_n = \arg \min_{\alpha \in \Theta} \varphi_n(\theta), \quad (6.7)$$

où

$$\varphi_n(\theta) = \frac{1}{n} \sum_{t=1}^n (X_t - f(X_{t-1}; \theta))^2. \quad (6.8)$$

Théorème 6.2.1. *Supposons que :*

1. *Sous \mathbb{P}_{θ_0} , la chaîne de Markov (X_t) est irréductible ;*
2. *pour un certain $k \geq 2$, $\mathbb{E}|\varepsilon_t|^k < +\infty$;*
3. *α_0 , la vraie valeur de α , appartient à l'intervalle $[0, 1[$;*
4. *l'espace des paramètres Θ , est un compact sous-ensemble de $[0, 1[\times]0, \infty[$.*

Alors, Le processus (X_t) est stationnaire et possède une unique mesure invariante, notée μ_{θ_0} telle que

$$\mu_{\theta_0}|\cdot|^k < \infty.$$

De plus, l'estimateur des moindres carrés $\hat{\theta}_n$ est fortement consistant, c.à.d.

$$\hat{\theta}_n \rightarrow \theta_0, \quad \mathbb{P}_{\theta_0} - p.s.$$

6.3 Propriétés du processus PRINAR(1)

Comme pour la classe RINAR, l'opérateur d'arrondi cause des difficultés pour l'étude théorique du processus PRINAR(1). Il est clair qu'un calcul explicite de la moyenne ou des coefficients d'autocorrélation n'est pas possible. Malgré tout, nous arrivons à borner ces derniers en utilisant les propriétés de la fonction arrondi et celles de la fonction partie fractionnaire.

Sur la moyenne :

Rappelons que, pour tout $a \in \mathbb{R}$, nous avons

$$\langle a \rangle = \lfloor a + \frac{1}{2} \rfloor \text{ et } \lfloor a \rfloor = a - \{a\},$$

où $\langle \cdot \rangle$ représente la fonction arrondi, $\lfloor \cdot \rfloor$ la fonction plancher "floor", et $\{ \cdot \}$ la partie fractionnaire.

De plus, nous trouvons

$$|a - \langle a \rangle| \leq \frac{1}{2} \text{ et } 0 \leq \{a\} < 1.$$

Dans la suite, nous supposons que les hypothèses du Théorème 6.2.1 sont vérifiées.

Notons que, avec les mêmes précédentes notations, le vrai processus PRINAR(1) peut être

écrit sous la forme suivante

$$\begin{aligned} X_t &= \langle \alpha_0 X_{t-1} \rangle + \varepsilon_t, \\ &= \lfloor \alpha_0 X_{t-1} + \frac{1}{2} \rfloor + \varepsilon_t, \\ &= \alpha_0 X_{t-1} + \frac{1}{2} - Z_{t-1} + \varepsilon_t, \end{aligned}$$

où $Z_t = \{\alpha_0 X_t + \frac{1}{2}\}$ et $\mathbb{E}\varepsilon_t = \mu_0$. Ici, les processus (X_t) et (Z_t) sont stationnaires, par conséquent

$$(1 - \alpha_0)\mathbb{E}X_t = \frac{1}{2} - \mathbb{E}Z_t + \mathbb{E}\varepsilon_t = \mu_0 + \frac{1}{2} - \mathbb{E}Z_t.$$

Comme $0 \leq Z_t < 1$ et $0 \leq \alpha_0 < 1$, nous obtenons

$$|\mathbb{E}X_t - \frac{\mu_0}{1 - \alpha_0}| = \frac{|\mathbb{E}Z_t - \frac{1}{2}|}{1 - \alpha_0} < \frac{\frac{1}{2}}{1 - \alpha_0}. \quad (6.9)$$

Rappelons que $\frac{\mu_0}{1 - \alpha_0}$ n'est autre que la moyenne d'un processus AR(1) réel, où α_0 est le coefficient de régression et μ_0 représente la moyenne du bruit associé.

Maintenant, posons $m = \mathbb{E}X_t$, $\forall t \in \mathbb{Z}$. Nous distinguons deux cas :

- Si le bruit est petit, $0 < \mu_0 < \frac{1}{2}$, alors

$$0 < m < \frac{\mu_0 + \frac{1}{2}}{1 - \alpha_0} < \frac{1}{1 - \alpha_0}. \quad (6.10)$$

Ici, la borne supérieur $\frac{1}{1 - \alpha_0}$ dépend juste de α_0 .

- Si le bruit est grand, $\mu_0 \gg 1$, alors

$$\frac{\mu_0 - \frac{1}{2}}{1 - \alpha_0} < m < \frac{\mu_0 + \frac{1}{2}}{1 - \alpha_0}. \quad (6.11)$$

Par conséquent,

$$m(1 - \alpha_0) - \frac{1}{2} < \mu_0 < m(1 - \alpha_0) + \frac{1}{2}. \quad (6.12)$$

De plus, si $\mu_0 \in \mathbb{N}^*$, nous obtenons

$$\mu_0 = \langle m(1 - \alpha_0) \rangle. \quad (6.13)$$

Sur le corrélogramme :

Comme $m = \mathbb{E}X_t \neq 0$, le calcul du coefficient d'autocorrélation du premier ordre du PRINAR(1) est compliqué. Soit le processus centré (Y_t) , où $Y_t = X_t - m$ qui a le même corrélogramme que (X_t) . Nous distinguons deux cas.

- Supposons que $\forall t \in \mathbb{Z}, \mathbb{E}X_t = m \in \mathbb{N}^*$. Il est clair que le processus $Y_t = X_t - m$ est à valeurs dans \mathbb{Z} et centré (c.à.d. $\mathbb{E}Y_t = 0$ et $|Y_t| \leq Y_t^2, \forall t \in \mathbb{Z}$). Ainsi, Nous avons

$$\begin{aligned} X_t &= \langle \alpha_0 X_{t-1} \rangle + \varepsilon_t, \\ X_t - m &= \langle \alpha_0 (X_{t-1} - m) + \alpha_0 m \rangle - m + \varepsilon_t. \end{aligned}$$

Alors,

$$Y_t = \langle \alpha_0 Y_{t-1} - c \rangle + \varepsilon_t, \text{ où } c = m(1 - \alpha_0) \text{ et } \mathbb{E}\varepsilon_t = \mu_0. \quad (6.14)$$

Par suite,

$$\text{cov}(X_t, X_{t+1}) = \text{cov}(Y_t, Y_{t+1}) = \alpha_0 V(Y_t) + \mathbb{E}Y_t(\langle \alpha_0 Y_t - c \rangle - (\alpha_0 Y_t - c)).$$

Comme $|a - \langle a \rangle| \leq \frac{1}{2}$, pour tout $a \in \mathbb{R}$, nous trouvons

$$|\rho(1) - \alpha_0| = \frac{|\mathbb{E}Y_t(\langle \alpha_0 Y_t - c \rangle - (\alpha_0 Y_t - c))|}{V(Y_t)} \leq \frac{1}{2} \frac{\mathbb{E}|Y_t|}{\mathbb{E}Y_t^2} \leq \frac{1}{2}. \quad (6.15)$$

- Supposons que $\forall t \in \mathbb{Z}, \mathbb{E}X_t = m \in \mathbb{R}^{+*}$. Ici, le processus centré (Y_t) , où $Y_t = X_t - m$, n'est plus à valeurs dans \mathbb{Z} (c.à.d. $|Y_t| \not\leq Y_t^2, \forall t \in \mathbb{Z}$) et nous pouvons l'écrire sous la forme suivante

$$Y_t = \langle \alpha_0 Y_{t-1} + \alpha_0 m \rangle + \mu_0 - m + \varepsilon'_t, \quad (6.16)$$

où $\varepsilon'_t = \varepsilon_t - \mu_0$. Ainsi, nous avons

$$|\rho(1) - \alpha_0| = \frac{|\mathbb{E}Y_t(\langle \alpha_0 Y_t + \alpha_0 m \rangle - (\alpha_0 Y_t + \alpha_0 m))|}{V(Y_t)} \leq \frac{1}{2} \frac{\mathbb{E}|Y_t|}{\mathbb{E}Y_t^2} = \frac{1}{2} \frac{\mathbb{E}|X_t - m|}{\mathbb{E}(X_t - m)^2}. \quad (6.17)$$

Rappelons que $[x]$ est la partie entière du nombre réel x . Il en découle,

$$\mathbb{E}|X_t - m| = \mu_{\theta_0}([m]) \varepsilon + \Sigma' = \Sigma' \left(1 + \frac{\mu_{\theta_0}([m]) \varepsilon}{\Sigma'} \right)$$

et

$$\mathbb{E}(X_t - m)^2 = \mu_{\theta_0}([m]) \varepsilon^2 + \Sigma,$$

où μ_{θ_0} est la mesure de probabilité stationnaire de la chaîne (X_t) ,

$$0 < \varepsilon = |[m] - m| < 1, \Sigma' = \sum_{i \neq [m]} \mu_{\theta_0}(i) |i - m|, \text{ et } \Sigma = \sum_{i \neq [m]} \mu_{\theta_0}(i) (i - m)^2.$$

Comme $\Sigma' \leq \Sigma \leq \mathbb{E}(X_t - m)^2$, nous obtenons

$$\frac{\mathbb{E}|X_t - m|}{\mathbb{E}(X_t - m)^2} \leq \left(1 + \frac{\mu_{\theta_0}([m])}{\Sigma'} \right).$$

Par conséquent,

$$|\rho(1) - \alpha_0| \leq \frac{1}{2} \left(1 + \frac{\mu_{\theta_0}([m])}{\Sigma'} \right). \quad (6.18)$$

6.3.1 Le processus PRINAR(1) recentré

Soient X_1, \dots, X_n des observations à valeurs entières positives modélisées par un processus PRINAR(1). Le procédé habituel de centrage de cette série chronologique semble inapproprié pour une modélisation à valeurs entières. En effet, comme $m \in \mathbb{R}^{+*}$, le processus centré $Y_t = X_t - m$ ne rentre pas dans le cadre des modèles considérés.

Nous distinguons deux cas :

- Supposons que $m \in \mathbb{N}^*$. Alors, le processus (Y_t) , défini par (6.14), est évidemment centré et à valeurs dans \mathbb{Z} . Ce dernier dépend de trois paramètres α_0 , $c = m(1 - \alpha_0)$ et μ_0 . De plus, il est clair qu'il ne possède pas la même représentation qu'un modèle RINAR(1). Par contre, ces coefficients d'autocorrélation ont les mêmes encadrements que ceux d'un modèle RINAR(1) de moyenne entière (par conséquent du processus RINAR(1) centré). En particulier, nous avons

$$|\rho(1) - \alpha_0| \leq \frac{1}{2}.$$

De plus, sous l'hypothèse $\mu_0 \in \mathbb{N}^*$ et d'après (6.5) (6.13) et (6.14), le processus centré (Y_t) s'écrit

$$Y_t = \langle \alpha_0 Y_{t-1} + c_1 \rangle + \varepsilon_t', \quad (6.19)$$

où $|c_1| = |\mu_0 - c| = |\langle c \rangle - c| \leq \frac{1}{2}$ et $\varepsilon_t' = \varepsilon_t - \mu_0$ (un bruit centré à valeurs dans \mathbb{Z}).

Le modèle présent (6.19) possède la même représentation qu'un processus RINAR(1).

- Maintenant, supposons que $m \in \mathbb{R}^{+*}$ et $m \geq \frac{1}{2}$ ($\langle m \rangle \geq 1$). Ici, nous proposons le processus (W_t) , où $W_t = X_t - \langle m \rangle$. Ce dernier peut être écrit sous la forme suivante

$$W_t = \langle \alpha_0 W_{t-1} - c' \rangle + \varepsilon_t, \text{ où } c' = \langle m \rangle(1 - \alpha_0) \text{ et } \mathbb{E}\varepsilon_t = \mu_0. \quad (6.20)$$

Il est clair que le processus (W_t) est à valeurs dans \mathbb{Z} mais il ne possède pas la même représentation qu'un modèle RINAR(1). D'autre part, sa moyenne est donnée par

$$|\mathbb{E}W_t| = |m - \langle m \rangle| \leq \frac{1}{2},$$

et ses coefficients d'autocorrélation possèdent les mêmes encadrements que ceux d'un RINAR(1) de moyenne non entière.

Par ailleurs, sous l'hypothèse $\mu_0 \in \mathbb{N}^*$ et d'après (6.5) (6.13) et (6.20), le processus (W_t) s'écrit

$$W_t = \langle \alpha W_t + c_1' \rangle + \varepsilon_t', \quad (6.21)$$

où $c'_1 = \mu_0 - c'$ et $\varepsilon'_t = \varepsilon_t - \mu$ (un bruit centré à valeurs dans \mathbb{Z}).

Ce dernier possède la même représentation qu'un modèle RINAR(1).

Notons que, comme $0 \leq \alpha_0 < 1$, nous trouvons

$$\begin{aligned} |c'_1| &= |\mu_0 - m(1 - \alpha) + m(1 - \alpha) - \langle m \rangle(1 - \alpha)| \\ &\leq |\mu_0 - m(1 - \alpha)| + |m(1 - \alpha) - \langle m \rangle(1 - \alpha)| \\ &\leq \frac{1}{2} + \frac{1}{2}(1 - \alpha) < 1. \end{aligned}$$

6.3.2 Étude de simulation

Au cours de notre étude nous constatons que le cas du modèle PRINAR(1) associé à un petit bruit représente beaucoup de curiosité.

Tout d'abord nous rappelons que, sous l'hypothèse $0 < \mu_0 < \frac{1}{2}$, nous avons

$$0 < m < \frac{1}{1 - \alpha_0} \text{ et } \begin{cases} |\rho(1) - \alpha_0| \leq \frac{1}{2}, & \text{si } m \in \mathbb{N}^* \\ |\rho(1) - \alpha_0| \leq \frac{1}{2} \left(1 + \frac{\mu_{\theta_0}([m])}{\Sigma'} \right), & \text{sinon.} \end{cases}$$

Pour mieux comprendre ce cas nous procédons par simulation. Soit α_0 et μ_0 les vraies valeurs des paramètres du modèle. Nous considérons une loi de Poisson de paramètre μ_0 pour le bruit (c.à.d. $\varepsilon_t \rightsquigarrow Po(\mu_0)$ et $\mathbb{E}\varepsilon_t = \mu_0$). Puis, nous simulons 1000 observations du processus PRINAR(1).

Nous proposons deux jeux de paramètres. pour le premier, nous associons au petit bruit un coefficient de régression (α_0) faible. Pour le deuxième, nous considérons $\alpha_0 > \frac{1}{2}$.

– $\alpha_0 = 0.25$ et $\mu_0 = 0.04$:

Les observations ci-dessous varient entre 0 et 1. Elles ressemblent à un exemple d'informations binaires. Le petit bruit et l'opérateur d'arrondi assurent une faible dépendance entre les observations. Pour cela, nous remarquons que les valeurs restent invariantes et égales à 0 sur longue plaque de temps, le passage à des valeurs supérieures dépend donc du bruit et de α_0 . La moyenne empirique est égale 0.039, l'écart-type est 0.193, et la médiane est 0. Notons que,

$$0 < \hat{m} = \frac{\sum_{i=1}^{1000} X_i}{1000} = 0.039 < \frac{1}{1 - \alpha_0} = \frac{4}{3}.$$

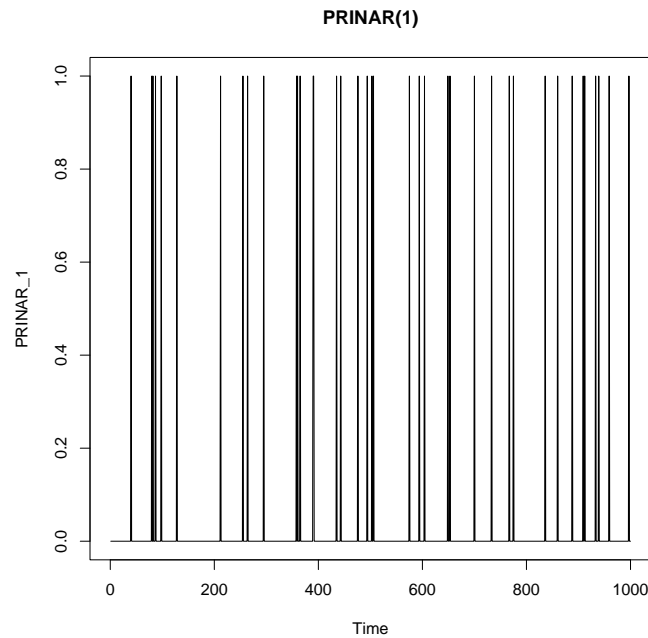


FIG. 6.1 – 1000 observations simulées du processus PRINAR(1) avec $\alpha_0 = 0.25$ et $\mu = 0.05$

Les figures suivantes donnent respectivement l'ACF et le PACF empiriques de la série simulée. Nous constatons qu'aucune autocorrélation n'est significative. Ainsi, si nous ne tenons pas en compte la nature entière des observations, cette série peut être interprétée comme un bruit blanc.

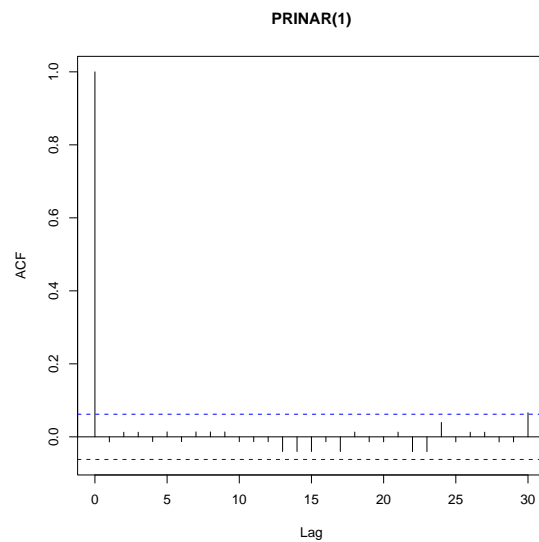


FIG. 6.2 – ACF de la série simulée

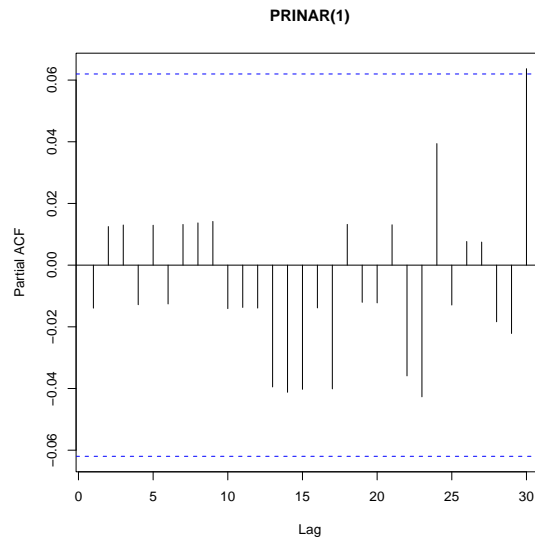
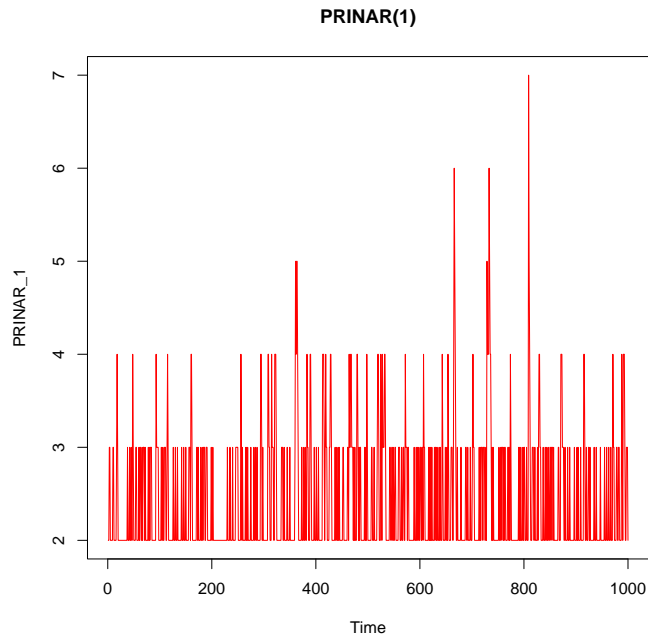


FIG. 6.3 – PACF de la série simulée

– $\alpha_0 = 0.78$ et $\mu_0 = 0.33$:

Les observations ci-dessous varient entre 2 et 7. La moyenne empirique est égale 2.442, l'écart-type est 0.679 et la médiane est 2. Notons que,

$$0 < \hat{m} = 2.442 < \frac{1}{1 - \alpha_0} = 4.54 \text{ et } \langle \hat{m} \rangle = \text{mediane} = 2.$$

FIG. 6.4 – 1000 observations simulées du processus PRINAR(1) avec $\alpha_0 = 0.78$ et $\mu = 0.33$

Rappelons que processus PRINAR(1) forme une chaîne de Markov homogène avec un espace d'états $E = \mathbb{N}$ et une probabilité de transition

$$\pi(x, y) = \mathbb{P} \{ \varepsilon_1 = y - \langle \alpha_0 x \rangle \}, \quad \forall x, y \in \mathbb{N}.$$

Ainsi, comme $\alpha_0 = 0.78$, nous avons

$$\mathbb{P} \{ X_{t+1} = 2 \mid X_t = 2 \} = \mathbb{P} \{ \varepsilon_1 = 0 \} = \exp(-\mu_0) \approx 0.72.$$

et

$$\mathbb{P} \{ X_{t+1} = 3 \mid X_t = 3 \} = \mathbb{P} \{ X_{t+1} = 3 \mid X_t = 2 \} \mathbb{P} \{ \varepsilon_1 = 1 \} = \exp(-\mu_0) \mu_0 \approx 0.24.$$

Les figures suivantes donnent respectivement l'ACF et le PACF empiriques de la série simulée.

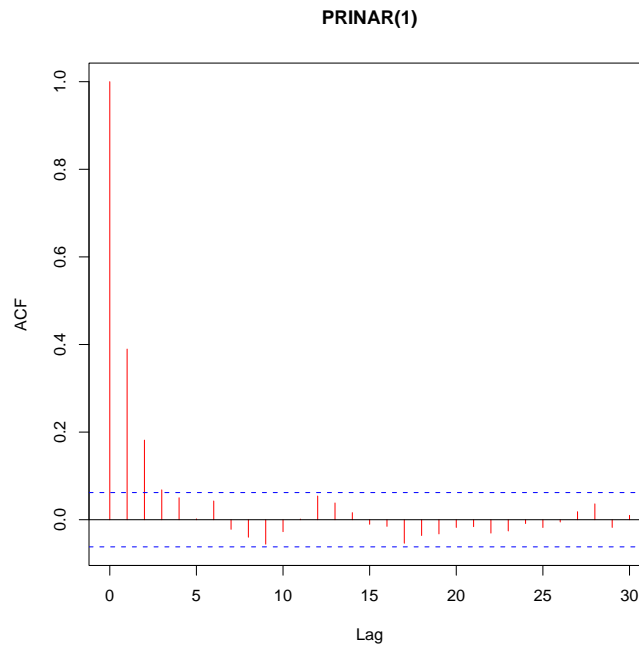


FIG. 6.5 – ACF de la série simulée

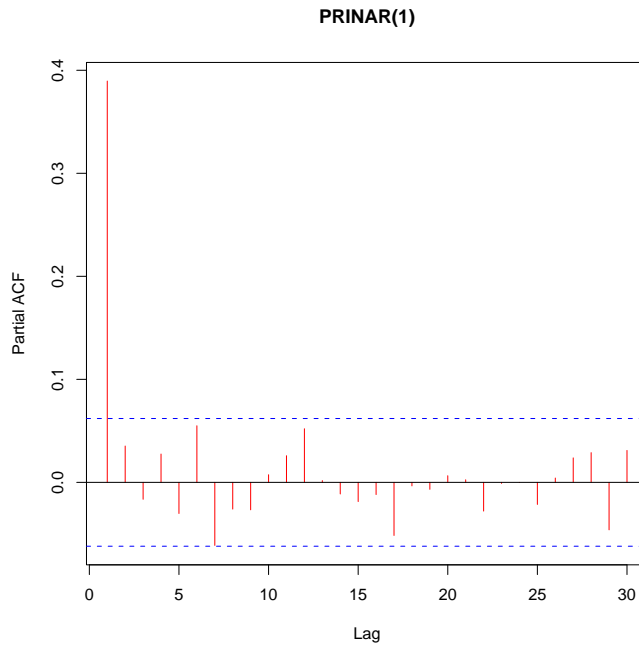


FIG. 6.6 – PACF de la série simulée

La décroissance géométrique des ACF empiriques et le fait que la fonction d'autocorrélation partielle du premier ordre soit l'unique valeur significative justifient le choix d'un processus AR(1) pour modéliser les observations. Ainsi, soit le modèle suivant

$$X_t = a_0 X_{t-1} + b_0 + \varepsilon_t.$$

Les estimateurs Yule-Walker des paramètres sont

$$\hat{a}_0 = \hat{\rho}(1) = 0.3895 \text{ et } \hat{b}_0 = \hat{m}(1 - \hat{a}_0) = 1.491.$$

Il est clair que $\theta_0 = (\hat{a}_0, \hat{b}_0)$ n'est pas une bonne estimation du paramètre $\theta = (\alpha, \mu)$ du modèle PRINAR(1).

D'autre part, rappelons que la prévision à un pas basée sur la fonction de régression, suivie par l'arrondi (pour revenir au support entier des observations) est donnée par

$$\hat{X}_{T+1} = \langle \hat{a}_0 X_T + \hat{b}_0 \rangle, \forall T \in \mathbb{Z}.$$

Soit t_0 un instant fixé. Nous remarquons que :

– Si $X_{t_0} = 2$ et $X_{t_0+1} = 2$, alors

$$\hat{X}_{t_0+1} = \langle 2 (0.3895) + 1.491 \rangle = 2 = X_{t_0+1};$$

- Si $X_{t_0} = 3$ et $X_{t_0+1} = 3$, alors

$$\hat{X}_{t_0+1} = \langle 3 (0.3895) + 1.491 \rangle = 3 = X_{t_0+1};$$

- Si $X_{t_0} = 2$ et $X_{t_0+1} = 3$ ou $X_{t_0} = 3$ et $X_{t_0+1} = 2$, alors

$$|X_{t_0+1} - \hat{X}_{t_0+1}| = 1.$$

Ainsi, nous définissons

$$\hat{X}_t = \langle \hat{a}_0 X_{t-1} + \hat{b}_0 \rangle, \quad \forall t = 2, \dots, 1000.$$

L'écart moyen absolue est défini par

$$EMA = \frac{\sum_{i=2}^{1000} |X_t - \hat{X}_t|}{999} = 0.461.$$

L'erreur quadratique moyenne est donnée par

$$EQM = \frac{\sum_{i=2}^{1000} |X_t - \hat{X}_t|^2}{999} = 0.559.$$

Donc, nous constatons que les performances de prévision du processus AR(1) sont plutôt bonnes.

6.4 Support de la loi stationnaire

Ici, nous admettons que les hypothèses du Théorème 6.2.1 sont vérifiées. Désormais, nous donnons les conditions nécessaires pour que la loi stationnaire, notée μ_{θ_0} , de la chaîne (X_t) possède un support fini.

Proposition 6.4.1. *Soit $A \in \mathbb{N}^*$ fixé. Supposons que $\text{supp}(\varepsilon_t) = [0, A]$. Alors, $\text{supp}(\mu_{\alpha_0}) \subseteq [0, B]$ où B est un entier tel que $B \geq \frac{1+2A}{2(1-\alpha_0)}$.*

Preuve. Rappelons que $\text{supp}(\varepsilon_t) = [0, A]$, c.à.d. $\varepsilon_t \leq A, \forall t \in \mathbb{Z}$. Notre objectif est de montrer que le support de μ_{θ_0} est fini. Tout d'abord, nous fixons $t_0 \in \mathbb{Z}$ et supposons que $X_{t_0} \leq B$ et $B \in \mathbb{N}$ tel que $B \geq \frac{1+2A}{2(1-\alpha_0)}$. Ainsi,

$$X_{t_0+1} = \langle \alpha_0 X_{t_0} \rangle + \varepsilon_{t_0+1} \leq \alpha_0 X_{t_0} + \frac{1}{2} + A \leq \alpha_0 B + \frac{1}{2} + A \leq B.$$

Par suite,

$$\forall t \in \mathbb{Z}, X_t \leq B \Rightarrow X_{t+1} \leq B.$$

Maintenant, nous posons $B_0 = \lceil \frac{1+2A}{2(1-\alpha_0)} \rceil$, où $\lceil x \rceil$ est le plus petit entier supérieur ou égale à x .

Ainsi, si $X_0 = B_0$, alors $\forall t \geq 1$ nous avons $X_t \leq B_0$. Soit μ_n la mesure empirique engendrée par X_1, \dots, X_n , définie par

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i}.$$

Le support de μ_n est $[0, B_0]$. Par conséquent, pour tout $c \notin \{0, \dots, B_0\}$,

$$\mu_{\theta_0}(c) = \lim_{n \rightarrow \infty} \mu_n(c) = 0.$$

□

Détaillons plus en avant la borne B de la proposition précédente.

– Si $\alpha_0 = 0$, alors $\forall t \in \mathbb{Z}$ nous avons $X_t = \varepsilon_t$. Par suite, nous obtenons $B = A + 1$.

– Si $0 < \alpha_0 < \frac{1}{2}$, alors $\frac{1+2A}{2(1-\alpha_0)} < 1 + 2A$. ainsi, il suffit de prendre $B = 1 + 2A$.

– Si $\frac{1}{2} \leq \alpha_0 < 1$, alors il existe $k(\alpha_0) \geq 2$ tel que $\alpha_0 \leq \frac{2k-1}{2k}$, ce qui implique que

$$\frac{1+2A}{2(1-\alpha_0)} \leq k(1+2A). \text{ Par conséquent, il suffit de prendre } B = k(1+2A).$$

Maintenant, nous considérons que le bruit ε_t suit une loi de Poisson de paramètre μ_0 (c.à.d. $\varepsilon_t \rightsquigarrow Po(\mu_0)$). Soit $a \in \mathbb{N}^*$, d'après l'inégalité de Bienaymé-Tchebychev, nous obtenons

$$\mathbb{P}(\mu_0 - a\sqrt{\mu_0} < \varepsilon_t < \mu_0 + a\sqrt{\mu_0}) \geq 1 - \frac{1}{a^2}.$$

Soit $a = 10$ et $0 < \mu_0 = 0.01 < \frac{1}{2}$ (le bruit est petit). Nous obtenons,

$$\mathbb{P}(\varepsilon_t \in \{0, 1\}) \geq 0.99.$$

Ici, nous constatons que dans ce cas le bruit est concentré sur l'intervalle $[0, A]$, où $A = 1$.

– Si $0 < \alpha_0 = 0.32 < \frac{1}{2}$, alors nous déduisons que (X_t) est plutôt concentrée sur l'intervalle $[0, B]$, où $B = 1 + 2A = 3$ ($\mathbb{P}(X_t \in \{0, 1, 2, 3\}) \geq 0.99$).

– Si $\alpha_0 = 0.78$, alors il existe $k(\alpha) = 3$ tel que $\alpha \leq \frac{2k-1}{2k}$. Ici, nous déduisons que la chaîne (X_t) est plutôt concentrée sur l'intervalle $[0, B]$, où $B = k(1+2A) = 9$ ($\mathbb{P}(X_t \in \{0, \dots, 9\}) \geq 0.99$).

Conclusion et perspectives

La modélisation des séries chronologiques à valeurs entières est un domaine scientifique assez récent. La plupart des modèles qui existent dans littérature sont appropriés pour analyser séries chronologiques à valeurs entières positives, en particulier les séries de comptage. Ces modèles sont basés sur les opérateurs d'amincissement et possèdent quelques limitations.

En revanche, l'utilisation de l'opérateur d'arrondi à l'entier le plus près pour générer des observations à valeurs entières paraît naturelle. Ainsi, comparée aux modèles autorégressifs basés sur les opérateurs d'amincissement, la classe RINAR introduite au cours de cette thèse possède plusieurs avantages. Plus précisément, les modèles RINAR ont une structure d'innovation simple, générée uniquement par le bruit. Par construction, la classe RINAR peut produire des séries chronologiques à valeurs entières ayant des observations négatives. D'autre part, les modèles RINAR peut produire des autocorrélations aussi riches que celles d'un AR(p) réel, y compris les autocorrélations négatives. De plus, la prévision à un pas, basée sur l'espérance conditionnelle (fonction de régression), est une valeur entière par construction des modèles.

Pour assurer la stationnarité et l'ergodicité du processus RINAR(1), nous supposons que la valeur absolue du coefficient de régression est strictement inférieur à 1. Cette hypothèse est similaire à celle qui garantit la stationnarité du modèle AR(1) réel. À cause de l'opérateur d'arrondi, l'identifiabilité du processus RINAR(1) a un comportement non standard. Ainsi, nous distinguons cas : Si la vraie valeur du coefficient de régression est irrationnelle, le problème d'identifiabilité n'apparaît pas. Par conséquent, l'estimateur des moindres carrés des paramètres est fortement consistant. Si la vraie valeur du coefficient de régression est rationnelle, alors le coefficient de régression est identifiable mais le paramètre de la constante l'est pas. Ainsi, l'estimateur des moindres carrés des paramètres converge presque sûrement vers une collection de points ayant le premier composant fixé

et égale à la vraie valeur du coefficient de régression, tandis que le deuxième composant varie dans un intervalle contenant la vraie valeur du paramètre de la constante et de taille proportionnelle au dénominateur de la vraie valeur du coefficient de régression. La taille maximale de cet intervalle est égale à $\frac{1}{2}$. Notons que cette fluctuation autour du paramètre de la constante n'a pas d'influence sur la prévision. Le modèle d'ordre supérieur RINAR(p), est une extension directe du processus RINAR(1). Nous mentionnons que dans le cas où tous les vraies valeurs des coefficients de régression, la taille maximale de l'intervalle de fluctuation du paramètre de la constante est inférieure à $\frac{1}{2}$. Par suite, le modèle vectoriel RINVAR est une généralisation du processus RINAR d'ordre supérieur. En effet, l'étude d'un processus RINVAR(1), dédié à la modélisation d'une série chronologique vectorielle de dimension p et à valeurs entières, revient à examiner p modèles RINAR(p) totalement séparés. Le modèle RINAR(1) centré est un cas particulier du processus RINAR(1), où le paramètre de la constante est nulle. Ce modèle ne possède pas de problème d'identifiabilité et par conséquent l'estimateur des moindres carrés de son paramètre est fortement consistant. Le modèle PRINAR(1) est consacré à l'analyse des séries chronologiques à valeurs entières positives. Comme son paramètre de la constante est hors de l'opérateur d'arrondi, le PRINAR(1) n'a pas un problème d'identifiabilité et par conséquent l'estimateur des moindres carrés des paramètres est fortement consistant.

Les modèles introduits au cours de cette thèse ont plusieurs atouts : une simple construction, des conditions de stationnarité et d'ergodicité plus générales que celles pour les modèles basés sur les opérateurs d'aminicissement, et des résultats satisfaisants concernant la consistance des estimateurs des moindres carrés. Cependant, à cause de l'opérateur d'arrondi, l'étude théoriques de certaines propriétés des modèles est compliquée. À titre d'exemple, pour un RINAR(1) stationnaire, nous n'arrivons pas à calculer explicitement la moyenne $m = \mathbb{E}X_t$. Toutefois, nous pouvons la trouver l'encadrement suivant,

$$|m - \frac{\lambda}{1 - \alpha}| \leq \frac{1}{2(1 - \alpha)},$$

où α est le coefficient de régression et λ est le paramètre de la constante. Rappelons que $\frac{\lambda}{1 - \alpha}$ n'est autre que la moyenne d'un modèle AR(1) réel ayant les mêmes paramètres. Par ailleurs, pour un processus RINAR(1) centré ($m = \mathbb{E}X_t = 0$) stationnaire, le calcul exact de ces coefficients d'autocorrélation $\rho(k)$ n'est pas possible. Cependant, pour $\rho(1)$ par exemple, nous trouvons l'encadrement suivant,

$$|\rho(1) - \alpha| \leq \frac{1}{2},$$

où α est le coefficient de régression du modèle $\text{RINAR}(1)$ centré. Notons que, pour un $\text{AR}(1)$ réel ayant le même coefficient de régression, nous avons $\rho(1) = \alpha$.

Soient X_1, \dots, X_n des observations à valeurs entières. Du point de vue pratique, une question naturelle se pose :

Comment nous validons le choix d'un processus RINAR pour modéliser ces observation ?

En premier lieu, nous traçons les coefficients d'autocorrélation empiriques simples (ACF) et partiels (PACF). Nous analysons ces coefficients comme si les observations sont à valeurs réelles. À titre d'exemple, nous examinons la décroissance des ACF et le nombre des coefficients PACF significatifs. Si par exemple, nous constatons que le processus $\text{AR}(1)$ est approprié, nous pouvons en déduire que le $\text{RINAR}(1)$ est bon candidat pour modéliser la série. Ainsi, nous calculons l'estimateur des moindres carrés en utilisant la méthode numérique proposée dans le Chapitre 2. Ensuite, nous vérifions si les résidus obtenus peuvent être interpréter comme un bruit blanc (aucune corrélation (ACF et PACF) des résidus n'est significative). Autrement, nous pouvons simuler des observations d'un processus $\text{RINAR}(1)$ ayant comme paramètres les valeurs calculées de l'estimateur des moindres carrés. Ensuite, nous traçons les ACF et les PACF de la série simulée et nous les comparons avec ceux de la série X_1, \dots, X_n .

Notons que, pour toutes les applications réelles introduites au cours de cette thèse, la valeur calculée de l'estimateur des moindres carrés du modèle $\text{RINAR}(p)$ par l'algorithme de la recherche dichotomique, est assez proche de celle de l'estimateur de Yule-Walker du modèle $\text{AR}(p)$ qui représente le point de départ de cet algorithme. Ceci explique les performances de prévision souvent analogues de ces modèles. Nous mentionnons qu'en général les valeurs de prévision d'un modèle $\text{AR}(p)$ sont réelles. Ainsi, nous considérons une opération d'arrondi finale pour retrouver le support entier des observations.

L'état actuel de nos travaux laisse entrevoir de nombreuses perspectives.

1. Améliorer certains résultats

- Pour tous les modèles proposés au cours de cette thèse nous supposons que le bruit (ε_t) est une suite de variables aléatoires i.i.d. à valeurs dans \mathbb{Z} . Une question se pose : Pouvons-nous affaiblir cette hypothèse ? En d'autres termes, il s'agit de trouver les conditions de stationnarité et d'ergodicité de ces modèles dans le cas où (ε_t) est un bruit blanc. Il est clair que sous cette hypothèse nous devons abandonner l'aspect Markovien.

- Pour assurer la stationnarité et l’ergodicité du modèle RINAR(p), nous imposons que la somme des valeurs absolues des coefficients de régression soit strictement inférieure à 1. Ceci implique que le rayon spectral de la matrice $|A|$ suivante soit strictement inférieur à 1,

$$|A| = \begin{pmatrix} |\alpha_1| \cdots |\alpha_p| \\ I_{p-1} & 0 \end{pmatrix},$$

avec I_{p-1} la matrice d’identité d’ordre $p-1$. Dorénavant, nous souhaitons étendre cette condition et retrouver celle qui garantit la stationnarité d’un AR(p) réel. Plus précisément, nous voulons que le rayon spectral de la matrice A soit strictement inférieur à 1. Par ailleurs, dans le cas du modèle RINVAR(1), nous visons aussi que le rayon spectral du coefficient de régression soit strictement inférieur à 1.

- Dans le cadre de l’étude des propriétés générales du modèle RINAR(1) centré stationnaire, nous trouvons les encadrements suivants pour ces coefficients d’auto-corrélation

$$|\rho(j) - \alpha^j| \leq \frac{1}{2} \sum_{i=0}^{j-1} |\alpha|^i, \quad \forall j \in \mathbb{N}^*.$$

À l’avenir, nous espérons trouver des bornes plus fines. Pour cela, nous pensons exploiter à titre d’exemple, la β -mélangeance du processus. Cette propriété n’a pas été annoncée au cours de nos travaux.

2. Développer quelques aspects inexploités

- L’étude de simulation dédié au modèle RINAR(1) centré (resp. PRINAR(1)), nous montre l’existence de deux cas dépendant du choix du bruit considéré. Ainsi, nous constatons qu’avec un grand bruit le coefficient d’autocorrélation empirique du premier ordre est assez proche du coefficient de régression. Cependant, quand le bruit est petit ces deux valeurs s’éloignent. De plus, les observations des séries simulées sont à faibles variations et restent souvent invariantes sur une longue plaque du temps. À titre d’exemple, avec un jeu de paramètre bien précis, le modèle PRINAR(1) produit des observations binaires. Ainsi, nous souhaitons expliquer théoriquement ces comportements.
- Pour assurer la positivité, la stationnarité et l’ergodicité du modèle PRINAR(1), nous supposons que son coefficient de régression soit strictement compris entre 0 et 1. Ainsi, nous visons clarifier la question suivante :
Est-ce que ce modèle peut produire des autocorrélations négatives ?

Notons que dans le cas d'une réponse insatisfaisante, le processus PRINAR(1) perd un des principaux avantages de la classe des modèles autorégressifs basés sur l'opérateur d'arrondi.

- Supposons que le support du bruit du modèle PRINAR(1) (resp. RINAR(1)) stationnaire est borné. Ainsi, nous obtenons que la loi stationnaire du processus l'est aussi. Ceci nécessite l'emploi d'autres techniques que celles proposées au cours de cette thèse, pour prouver la consistance de l'estimateur des moindres carrés.
- L'étude du modèle RINVAR(1), dédié à l'analyse d'une série chronologique vectorielle de dimension p à valeurs entières, revient à traiter p modèles RINAR(p) séparés. Ceci grâce à la définition de l'opérateur arrondi et au fait que tous les composants ont le même support. À l'avenir, nous souhaitons examiner le cas où un au moins de ces composants a un support borné.

3. Étudier la vitesse de convergence de l'estimateur des moindres carrés

Afin de compléter notre étude dédiée au processus RINAR(1), notre intérêt s'est tourné vers l'étude de la vitesse de convergence de l'estimateur des moindres carrés, considéré pour estimer les paramètres du modèle. Ce travail, mené en collaboration avec Saumard, A.¹, est en cours de préparation. Dans la suite, nous citons quelques uns des arguments principaux utilisés pour aboutir à un tel résultat.

Rappelons que, le modèle RINAR(1) stationnaire est défini par

$$X_t = \langle \alpha X_{t-1} + \lambda \rangle + \varepsilon_t = f(X_{t-1}; \theta) + \varepsilon_t,$$

où $\theta = (\alpha, \lambda) \in \Theta$ l'espace des paramètres compact et sous-ensemble de $] -1, 1[\times \mathbb{R}$. Soit $\theta_0 = (\alpha_0, \lambda_0)$ la vraie valeur des paramètres du modèle, μ la mesure invariante de ce processus (sous θ_0), et $\mathcal{G} = (g_\theta)_{\theta \in \Theta}$, où $g_\theta((x, y)) = (y - f(x; \theta))^2$, $(x, y) \in \mathbb{Z}^2$, la classe des fonctions de contraste. L'estimateur des moindres carrés $\hat{\theta}_n$ peut être présenté de la manière suivante

$$g_{\hat{\theta}_n} = \arg \min_{g_\theta \in \mathcal{G}} P_n g_\theta, \quad (*)$$

où P_n est la mesure empirique associée à la chaîne double $Y_t = (X_{t-1}, X_t)$ ayant P comme mesure invariante (sous θ_0). Notons que, la définition (*) nous donne $P_n(g_{\hat{\theta}_n} - g_{\theta_0}) \leq 0$. Comme les fonctions de contraste sont irrégulières, les techniques usuelles basées sur la dérivabilité de ces fonctions sont désormais inadéquates.

¹Adrien Saumard : IRMAR, université de Rennes 1.

Ainsi, pour contrôler la vitesse de convergence nous utilisons des méthodes non paramétriques basées sur les processus empiriques.

Tout d'abord, nous montrons que le processus RINAR(1) est β -mélangeant (voir, Mokkadem, A.² (1990)) et que \mathcal{G} est une classe de sous-graphes de Vapnik-Chervonenkis (VC-subgraph class), voir par exemple, van der Vaart, A. W. & Wellner, J. A.³ (1996) et Dudley, R. M.⁴ (1999)). Ensuite, nous utilisons une technique de contrôle du supremum du processus empirique par entropie universelle pour les données mélangeantes, développée par Doukhan, P., Massart, P. & Rio, E.⁵ (1994) puis Rio, E.⁶ (1998).

Finalement, grâce à cette technique nous contrôlons à distance finie la taille de l'espérance du supremum du processus empirique sur la classe \mathcal{G} , et donc le risque quadratique moyen de la série estimée à l'instant n (le modèle sous $\hat{\theta}_n$) notée \hat{X}_n , par rapport à la vraie série chronologique inconnue (le modèle sous θ_0), notée X_0 .

En effet, le lien entre ces deux quantités est donné par le simple calcul suivant :

$$\begin{aligned} \mathbb{E} \left[\mu(\hat{X}_n - X_0)^2 \right] &= \mathbb{E} \left[\mu(f(\cdot; \hat{\theta}_n) - f(\cdot; \theta_0))^2 \right], \\ &= \mathbb{E} \left[P(g_{\hat{\theta}_n} - g_{\theta_0}) \right], \\ &\leq \mathbb{E} \left[(P - P_n)(g_{\hat{\theta}_n} - g_{\theta_0}) \right], \quad \text{d'après l'équation } (*), \\ &\leq \frac{1}{\sqrt{n}} \mathbb{E} \left[\sup_{g_{\theta} \in \mathcal{G}} |\alpha_n(g_{\theta} - g_{\theta_0})| \right], \end{aligned}$$

où $\alpha_n = \sqrt{n}(P - P_n)$.

²Propriétés de mélange des processus autorégressifs polynomiaux. Ann. Inst. H. Poincaré Probab. Statist. v. 26, no. 2, p. 219–260.

³Weak convergence and empirical processes. With applications to statistics, Springer-Verlag, New York.

⁴Uniform central limit theorems, Cambridge University Press.

⁵The functional central limit theorem for strongly mixing processes. Ann. Inst. H. Poincaré Probab. Statist. v. 30, no. 1, p. 63–82.

⁶Processus empiriques absolument réguliers et entropie universelle. Probab. Theory Related Fields v. 111, no. 4, p. 585–608.

4. Introduire d'autres modèles

La classe des modèles RINAR est inspirée des modèles autorégressifs réels. En effet, l'opérateur d'arrondi est considéré comme une censure sur la fonction de régression linéaire d'un modèle AR. Dorénavant, nous proposons d'appliquer la même philosophie à la classe des modèles ARCH/GARCH. En réalité, les séries financières de variation et à valeurs entières, comme dans le cas à valeurs réelles, forment une suite de bruit blanc. Les modèles du type autorégressifs introduits au cours de cette thèse deviennent ainsi inadaptés. Il est alors naturel d'adapter la construction ARCH/GARCH au contexte de valeurs entières.

Table des figures

2.1	La fonction arrondi	21
2.2	La fonction partie fractionnaire	21
2.3	La fonction partie entière	22
2.4	La subdivision de l'intervalle $[0, 1[$ en 4 intervalles de taille $\frac{1}{4}$	44
2.5	La subdivision de l'intervalle $[0, 1[$ en 5 intervalles de taille $\frac{1}{5}$	44
2.6	La nouvelle subdivision de l'intervalle $[0, 1[$, avec $q = 7$	45
2.7	Le passage de l'étape i à $i + 1$ pour la recherche dichotomique.	58
2.8	L'histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = -0.4$	62
2.9	L'histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = 1.44$	62
2.10	L'histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = -0.2$	63
2.11	L'histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = -2.96$	63
2.12	L'histogramme des $(\hat{\alpha}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_0 = 0.3$	64
2.13	L'histogramme des $(\hat{\lambda}_{n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda_0 = 5.65$	64
2.14	70 observations provenant des résultats consécutifs d'un processus chimique, source : O'Donovan.	65
2.15	ACF des 70 observations d'O'Donovan.	66
2.16	PACF des 70 observations d'O'Donovan.	67
2.17	Test sur les résidues obtenus en utilisant AR(1) pour modéliser les données d'O'Donovan.	68
2.18	Les résidus obtenus en utilisant le processus RINAR(1).	69
2.19	ACF des résidus.	70
2.20	PACF des résidus.	70
2.21	Les taux de variation annuels de la population Suédoise, 1750 – 1849.	72
2.22	ACF et PACF des taux de variation annuels de la population Suédoise.	73
2.23	Les résidus obtenus en utilisant le processus RINAR(1).	75

2.24	ACF et PACF des résidus.	75
2.25	Les 20 dernières observations des taux de variation de la population Suédoise et leurs prévisions basées sur les modèles RINAR(1) et AR(1).	77
3.1	L'histogramme des $(\hat{\alpha}_{1n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_1^* = 0.12$	107
3.2	L'histogramme des $(\hat{\lambda}_{2n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_2^* = 0.375$	108
3.3	L'histogramme des $(\hat{\lambda}_{3n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_3^* = 0.2$	108
3.4	L'histogramme des $(\hat{\lambda}_{4n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\alpha_4^* = -0.25$	109
3.5	L'histogramme des $(\hat{\lambda}_{4n,i})_{1 \leq i \leq 500}$ avec $n = 500$ et $\lambda^* = 2.5$	109
3.6	Les données de Fürth.	110
3.7	L'ACF et le PACF de la série Fürth.	111
3.8	Les erreurs de prévision $\hat{\varepsilon}_{T+1} = X_{T+1} - \hat{X}_{T+1}$, $400 \leq T \leq 504$, provenant du modèle RINAR(2).	113
3.9	ACF empirique de la série simulée du modèle RINAR(2), avec $\theta_0 = \hat{\theta}_n =$ $(0.818, -0.23, 0.697)$	114
3.10	PACF empirique de la série simulée du modèle RINAR(2), avec $\theta_0 = \hat{\theta}_n =$ $(0.818, -0.23, 0.697)$	115
4.1	Les indices de récolte annuels de la population Suédoise, h_t , 1750 – 1849.	138
4.2	ACF et le PACF empirique de la série h_t	139
4.3	Les taux de variation de la population Suédoise et leurs prévisions basées sur le modèle de McCleary & Hay et celles basées sur le modèle RINVAR(1).	143
5.1	La fonction "floor"	146
5.2	La fonction partie fractionnaire	147
5.3	La fonction arrondi	148
5.4	Les observations simulées du RINAR(1) centré, avec $\alpha_0 = 0.83$ et $\mu = 9.453$	156
5.5	ACF de la série simulée	157
5.6	ACF théorique d'un AR(1) avec $\alpha_0 = 0.83$	157
5.7	Les observations simulées du RINAR(1) centré, avec $\alpha_0 = 0.83$ et $\mu = 0.1274$	158
5.8	ACF de la série simulée	159
5.9	Histogramme des estimations de $\alpha_0 = 0.83$ (avec grand bruit)	161
5.10	Histogramme des $\hat{\rho}_k(1)$, $1 \leq k \leq 500$	162
5.11	Histogramme des estimations de $\alpha_0 = 0.83$ (avec petit bruit)	162

6.1	1000 observations simulées du processus PRINAR(1) avec $\alpha_0 = 0.25$ et $\mu = 0.05$	173
6.2	ACF de la série simulée	173
6.3	PACF de la série simulée	174
6.4	1000 observations simulées du processus PRINAR(1) avec $\alpha_0 = 0.78$ et $\mu = 0.33$	174
6.5	ACF de la série simulée	175
6.6	PACF de la série simulée	176

Bibliographie

- [1] M. A. Al-Osh and E. E. A. A. Aly. First order autoregressive time series with negative binomial and geometric marginals. Comm. Statist. Theory Methods, 21(9) :2483–2492, 1992.
- [2] M. A. Al-Osh and A. A. Alzaid. First-order integer-valued autoregressive (INAR(1)) process. J. Time Ser. Anal., 8(3) :261–275, 1987.
- [3] M. A. Al-Osh and A. A. Alzaid. Integer-valued moving average (INMA) process. Statist. Hefte, 29, 1988.
- [4] M. A. Al-Osh and A. A. Alzaid. Binomial autoregressive moving average models. Comm. Statist. Stochastic Models, 7(2) :261–282, 1991.
- [5] A. A. Alzaid and M. A. Al-Osh. First-order integer-valued autoregressive (INAR (1)) process : distributional and regression properties. Statist. Neerlandica, 42(1) :53–61, 1988.
- [6] A. A. Alzaid and M. A. Al-Osh. An integer-valued p th-order autoregressive structure (INAR(p)) process. J. Appl. Probab., 27(2) :314–324, 1990.
- [7] A. A. Alzaid and M. A. Al-Osh. Some autoregressive moving average processes with generalized Poisson marginal distributions. Ann. Inst. Statist. Math., 45(2) :223–232, 1993.
- [8] R. S. Ambagaspitiya and N. Balakrishnan. On the compound generalized poisson distributions. ASTIN Bulletin, 24(2) :255–263, 1994.
- [9] K. Brännäs and A. Hall. Estimation in integer-valued moving average models. Appl. Stoch. Models Bus. Ind., 17(3) :277–291, 2001.
- [10] K. Brännäs and J. Hellström. Generalized integer-valued autoregression. Econometric Rev., 20(4) :425–443, 2001.

- [11] K. Brännäs, Hellström J., and Nordström J. A new approach to modelling and forecasting monthly guest nights in hotels. International Journal of Forecasting, 18(1) :19–30, 2002.
- [12] P. J. Brockwell and R.A. Davis. Introduction to time series and forecasting. Springer Texts in Statistics. Springer-Verlag, New York, second edition, 2002.
- [13] P. C. Consul. Generalized Poisson distributions, volume 99. Dekker, New York, 1989.
- [14] I. M. M. Da Silva. Contributions to the analysis of discrete-valued time series. PhD thesis, University of Porto, Portugal, 2005.
- [15] I. M. M. Da Silva and M. E. Da Silva. Asymptotic distribution of the Yule-Walker estimator for INAR(p) processes. Statist. Probab. Lett., 76(15) :1655–1663, 2006.
- [16] M. E. Da Silva and V. L. Oliveira. Difference equations for the higher-order moments and cumulants of the Inar(1) model. J. Time Ser. Anal., 25(3) :317–333, 2004.
- [17] M. E. Da Silva and V. L. Oliveira. Difference equations for the higher order moments and cumulants of the INAR(p) model. J. Time Ser. Anal., 26(1) :17–36, 2005.
- [18] J.P. Dion, G. Gauthier, and A. Latour. Branching processes with immigration and integer-valued time series. Serdica Math. J., 21(2) :123–136, 1995.
- [19] P. Doukhan, A. Latour, and D. Oraichi. A simple integer-valued bilinear times series models. Advances in Applied Probability, 38 :1–20, 2006.
- [20] F. C. Drost, R. van den Akker, and B. J. M. Werker. Local asymptotic normality and efficient estimation for INAR(P) models. J. Time Ser. Anal., 29(5) :783–801, 2008.
- [21] J.G. Du and Y. Li. The integer-valued autoregressive (INAR(p)) model. J. Times Ser. Anal., 12 :129–142, 1991.
- [22] M. Duflo. Random Iterative Models, volume 34 of Applications of Mathematics (New York). Springer-Verlag, Berlin, 1997.
- [23] W. Feller. An introduction to probability theory and its applications. Vol. I. Third edition. Wiley, New York, 1968.
- [24] R. Ferland, A. Latour, and D. Oraichi. Integer-valued GARCH process. J. of Time Ser. Anal., 27(6) :923–942, 2006.
- [25] R. K. Freeland. Statistical analysis of discrete time series with applications to the analysis of workers compensation claims data. PhD thesis, University of British Columbia, Canada, 1998.

- [26] R. K. Freeland and B. P. M. McCabe. Analysis of low count time series data by Poisson autoregression. J. Time Ser. Anal., 25(5) :701–722, 2004.
- [27] R. K. Freeland and B. P. M. McCabe. Asymptotic properties of CLS estimators in the Poisson AR(1) model. Statist. Probab. Lett., 73(2) :147–153, 2005.
- [28] R. Fürth. Statistic und wahrscheinlichkeitsnachwirkung. Physikalische Zeitschrift, 19 :421–426, 1918.
- [29] G. Gauthier. Modèle de type autorégressif pour les séries chronologiques à valeurs entières non négatives. Master’s thesis, Département de mathématiques et d’informatique, 1991. 179 pages.
- [30] G. Gauthier and A. Latour. Convergence forte des estimateurs des paramètres d’un processus GENAR(p). Ann. Sci. Math. Québec, 18(1) :49–71, 1994.
- [31] C. C. Heyde and E. Seneta. Estimation theory for growth and immigration rates in a multiplicative process. J. Appl. Probability, 9 :235–256, 1972.
- [32] P. A. Jacobs and P. A. W. Lewis. Discrete time series generated by mixtures. I. Correlational and runs properties. J. Roy. Statist. Soc. Ser. B, 40(1) :94–105, 1978.
- [33] P. A. Jacobs and P. A. W. Lewis. Discrete time series generated by mixtures. II. Asymptotic properties. J. Roy. Statist. Soc. Ser. B, 40(2) :222–228, 1978.
- [34] P. A. Jacobs and P. A. W. Lewis. Stationary discrete autoregressive-moving average time series generated by mixtures. J. Time Ser. Anal., 4(1) :19–36, 1983.
- [35] H. Joe. Time series models with univariate margins in the convolution-closed infinitely divisible class. J. Appl. Probab., 33(3) :664–677, 1996.
- [36] P. Johansson. Speed limitation and motorway causalities : a time series count data regression approach. Accident Anal. Prev., 28 :73–87, 1996.
- [37] R. C. Jung, M. Kukuk, and R. Liesenfeld. Time series of count data : modeling, estimation and diagnostics. Comput. Statist. Data Anal., 51(4) :2350–2364, 2006.
- [38] R. C. Jung, G. Ronning, and A. R. Tremayne. Estimation in conditional first order autoregression with discrete support. Statist. Papers, 46(2) :195–224, 2005.
- [39] R. C. Jung and A. R. Tremayne. Testing for serial dependence in time series models of counts. J. Time Ser. Anal., 24(1) :65–84, 2003.
- [40] R. C. Jung and A. R. Tremayne. Binomial thinning models for integer time series. Stat. Model., 6(2) :81–96, 2006.

- [41] M. Kachour. First-order rounded integer-valued vectorial autoregressive (RINVAR(1)) process. Preprint, University of Rennes 1, 2008-2009.
- [42] M. Kachour. p -order rounded integer-valued autoregressive (RINAR(p)) process. Preprint, University of Rennes 1, 2008-2009.
- [43] M. Kachour and J.F. Yao. The first-order rounded integer-valued autoregressive (RINAR(1)) process. J. Time Ser. Anal., 4 :417–448, July 2009.
- [44] B. Kedem and K. Fokianos. Regression models for time series analysis. Wiley Series in Probability and Statistics. Hoboken, NJ, 2002.
- [45] D. Lambert and C. Liu. Adaptive thresholds : Monitoring streams of network counts. J. Am. Stat. Assoc., 101 :78–88, 2006.
- [46] A. Latour. The multivariate GINAR(p) process. Adv. in Appl. Probab., 29(1) :228–248, 1997.
- [47] A. Latour. Existence and stochastic structure of a non-negative integer-valued autoregressive process. J. Time Ser. Anal., 19(4) :439–455, 1998.
- [48] E. Lukacs. Characteristic functions. Griffin, London, second edition, 1970.
- [49] R. McCleary and R.A. Hay. Applied Time Series Analysis for the Social Sciences. Sage Publications, 1980.
- [50] E. McKenzie. Some simple models for discrete variate time series. Water Resources Bulletin, 21(4) :645–650, 1985.
- [51] E. McKenzie. Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. Adv. in Appl. Probab., 18(3) :679–705, 1986.
- [52] E McKenzie. Some ARMA models for dependent sequences of Poisson counts. Adv. in Appl. Probab., 20 :822–835, 1988.
- [53] Eddie McKenzie. Discrete variate time series. In Stochastic processes : modelling and simulation, volume 21 of Handbook of Statist., pages 573–606. Amsterdam, 2003.
- [54] T. M. Mills and E. Seneta. Goodness-of-fit for a branching process with immigration using sample partial autocorrelations. Stochastic Process. Appl., 33(1), 1989.
- [55] T.M. O'Donovan. Short Term Forecasting : An Introduction to the Box-Jenkins Approach. Wiley, 1983.
- [56] F. Pokropp, W. Seidel, A. Begun, M. Heidenereich, and K. Sever. Control charts for the number of children injured in traffic accident. Frontiers in Stat. Qual. Control, 8 :151–171, 2006.

- [57] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical recipes. Cambridge University Press, 2007.
- [58] M. M. Ristić, H. S. Bakouch, and A. S. Nastić. A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. J. Statist. Plann. Inference, 139(7) :2218–2226, 2009.
- [59] F. W. Steutel and K. van Harn. Discrete analogues of self-decomposability and stability. Ann. Probab., 7(5) :893–899, 1979.
- [60] D.S. Thomas. Social and Economic Aspects of Swedish Population Movements, 1750-1933. Macmillan, New York, 1940.
- [61] A. W. van der Vaart. Asymptotic Statistics. Cambridge University Press, 1998.
- [62] C. H. Weiß. Controlling correlated processes of poisson counts. Qual. Reliab. Engng. Int., 23(6) :741–754, 2007.
- [63] C. H. Weiß. Serial dependence and regression of Poisson INARMA models. J. Statist. Plann. Inference, 138(10) :2975–2990, 2008.
- [64] C. H. Weiß. Thinning operations for modeling time series of counts—a survey. AStA Adv. Stat. Anal., 92(3) :319–341, 2008.
- [65] N. Ye, J. Giordano, and J. Feldman. A process control approach to cyber attack detection. Commun. ACM, 44(8) :76–82, 2001.
- [66] H. Zheng, I. V. Basawa, and S. Datta. First-order random coefficient integer-valued autoregressive processes. J. Statist. Plann. Inference, 137(1) :212–229, 2007.
- [67] R. Zhu and H. Joe. A new type of discrete self-decomposability and its application to continuous-time Markov processes for modeling count data time series. Stoch. Models, 19(2) :235–254, 2003.
- [68] R. Zhu and H. Joe. Modelling count data time series with Markov processes based on binomial thinning. J. Time Ser. Anal., 27(5) :725–738, 2006.

VU :

VU :

Le Directeur de Thèse

Le Responsable de l'École Doctorale

VU pour autorisation de soutenance

Rennes, le

Le Président de l'Université de Rennes 1

Guy CATHELINEAU

VU après soutenance pour autorisation de publication :

Le Président de Jury,

Résumé :

Dans certaines situations il devient nécessaire de traiter les séries chronologiques à valeurs entières. Au premier regard, l'analyse de telle série peut présenter quelques difficultés, notamment si l'analyse est basée sur quelques modèles stochastiques. Ces modèles doivent refléter la particularité entière de la série observée. De nombreuses tentatives ont été faites pour définir des modèles qui peuvent être utilisés pour décrire les séries chronologiques à valeurs entières. La plupart des modèles proposés sont basés sur l'opérateur d'amincissement et possèdent les mêmes propriétés que les modèles à valeurs réelles bien-connus dans la littérature.

L'objectif de cette thèse est d'étudier les modèles autorégressifs à valeurs entières. Nous introduisons une nouvelle classe de modèles basés sur l'opérateur d'arrondi. Par rapport aux modèles existants, la nouvelle classe a plusieurs avantages : structure d'innovation simple, coefficients de régression avec des signes arbitraires, valeurs négatives possibles pour la série chronologiques et pour la fonction d'autocorrélation. Nous étudions la stationnarité des modèles et la consistance forte de l'estimateur des moindres carrés proposé pour estimer les paramètres. Nous analysons quelques séries chronologiques à valeurs entières bien-connues avec les modèles introduits.

Mots clefs : Séries chronologiques à valeurs entières, opérateur d'amincissement, opérateur d'arrondi, identifiabilité, estimateur des moindres carrés, consistance forte.

Abstract :

In many practical situations we deal with integer-valued time series. The analysis of such a time series present some difficulties, namely where the analysis is based on some stochastic models. These models must reflect the integer peculiarity of the observed series. Many attempts have been made to define some models which can be used to describe integer-valued time series. Most of the proposed models are based on the thinning operator and they have the same properties as the real-valued models well-known in the literature.

The aim of this thesis is to study the integer-valued autoregressive models. We introduce a new class of models based on the rounding operator. Compared to the existent models, the new class has several advantages : simple innovation structure, autoregressive coefficients with arbitrary signs, possible negative values for time series and for the autocorrelation function.

We study the stationarity of the models and the strong consistency of the least squares estimator proposed to estimate the parameters. We analyze some well-known time series with the introduced models.

Key words : Integer-valued time series, thinning operator, rounding operator, identifiability, least squares estimator, strong consistency.